

А. Т. Фоменко

ТОПОЛОГИЧЕСКИЕ ВАРИАЦИОННЫЕ ЗАДАЧИ

ИЗДАТЕЛЬСТВО
МОСКОВСКОГО
УНИВЕРСИТЕТА



А. Т. Фоменко

ТОПОЛОГИЧЕСКИЕ ВАРИАЦИОННЫЕ ЗАДАЧИ

ИЗДАТЕЛЬСТВО
МОСКОВСКОГО УНИВЕРСИТЕТА
1984

Фоменко А. Т. Топологические вариационные задачи. — М.: Изд-во Моск. ун-та, 1984. — 216 с.

В книге с единой геометрической точки зрения излагаются следующие основные темы: гомологии и когомологии, теория препятствий, современные аспекты теории Морса на гладких многообразиях, топология многообразий малой размерности, последние достижения, связанные с известной проблемой А. Пуанкаре, теория минимальных поверхностей и гармонических отображений. Излагается также экспериментальный материал, лежащий в основе многих современных понятий многомерного вариационного исчисления, в частности, описываются физические опыты с поверхностями раздела двух сред. Многие из этих вопросов излагаются в литературе учебного характера впервые. Книга рассчитана на широкие круги читателей, интересующихся современными методами геометрии, топологии и их приложениями.

Библиогр. 63 назв. Ил. 187.

Заставки к главам и рисунки в тексте выполнены А. Т. Фоменко

Рецензенты:

проф. В. В. Федорчук;
доктор физ.-матем. наук В. В. Козлов

Печатается по постановлению
Редакционно-издательского совета
Московского университета

Анатолий Тимофеевич Фоменко

ТОПОЛОГИЧЕСКИЕ ВАРИАЦИОННЫЕ ЗАДАЧИ

Зав. редакцией С. И. Зеленский. Редактор А. А. Локшин. Обложка художника Н. Н. Сенько. Художественный редактор Л. В. Мухина. Технический редактор К. С. Чистякова. Корректоры В. П. Кададинская Н. И. Коновалова.

ИБ № 1935
КБ № 37—27—84

Сдано в набор 29.02.84 Подписано к печати 12.10.84 Л-79879
Формат 60×90/16 Бумага тип. № 3 Гарнитура литературная
Высокая печать. Усл. печ. л. 13,5 Уч.-изд. л. 14,04 Тираж 3600 экз.
Заказ 344 Цена 45 коп. Изд. № 3185

Ордена «Знак Почета» издательство Московского университета.
103009, Москва, ул. Герцена, 5/7. Типография ордена «Знак Почета»
изд-ва МГУ. 119899, Москва, Ленинские горы.

Ф 1702040000—096 КБ № 37—27—84
077(02)—84

© Издательство Московского университета, 1984 г.

Предисловие	5
Глава 1. НЕКОТОРЫЕ ЭЛЕМЕНТАРНЫЕ ТОПОЛОГИЧЕСКИЕ СВЕДЕНИЯ	9
§ 1. Группы сингулярных и клеточных гомологий	10
1. Сингулярные цепи и группы гомологий. — 2. Клеточные комплексы, барицентрические подразделения. — 3. Клеточные гомологии и вычисление сингулярных гомологий сферы. — 4. Теорема о совпадении сингулярных и клеточных гомологий конечного клеточного комплекса. — 5. Геометрическое опре- деление групп клеточных гомологий. — 6. Простейшие при- меры вычисления групп клеточных гомологий.	
§ 2. Группы когомологий и препятствия к продолжению отобра- жений	25
1. Сингулярные коцепи и оператор δ . — 2. Задача продолже- ния непрерывного отображения с подпространства на все про- странство. — 3. Препятствие к продолжению отображения. — 4. Случай существования ретракции пространства на подпро- странство, гомеоморфное сфере.	
§ 3. Расслоения	36
1. Определение локально тривиального расслоения. — 2. При- меры расслоений. — 3. Геометрия расслоения Хопфа. — 4. Геометрия расслоения единичных касательных векторов к сфере и индекс векторного поля.	
Глава 2. ФУНКЦИИ НА МНОГООБРАЗИЯХ	47
§ 1. Точные функции Морса	48
§ 2. Многозначный аналог теории Морса	53
Глава 3. МНОГООБРАЗИЯ МАЛЫХ РАЗМЕРНОСТЕЙ	61
§ 1. Гомеоморфизмы двумерных поверхностей	62
1. Группа гомеотопий и классификация трехмерных многооб- разий рода 1. — 2. Система образующих группы гомеотопий полного кренделя.	
§ 2. Алгоритм распознавания стандартной трехмерной сферы в классе диаграмм Хегора рода два	68
1. Графы Уайтхеда, кодирующие трехмерные многообразия. Операции индексов один и два. — 2. Волны, разбивающие вершины, и операция упрощения графов Уайтхеда. — 3. Су- ществование разбивающей вершины на графах Уайтхеда рода два, соответствующих стандартной трехмерной сфере. Алго- ритм распознавания сферы.	
§ 3. О решении четырехмерной проблемы Пуанкаре: любая гомото- пическая четырехмерная сфера гомеоморфна стандартной сфере	80

1. Физические опыты Плато и способы получения мыльных пленок. — 2. Физические принципы, лежащие в основе образования мыльных пленок. — 3. Экстремальные свойства мыльных пленок и минимальность их площади. Свойства границы раздела двух сред. — 4. Граница раздела двух сред, находящихся в равновесии, является поверхностью постоянной средней кривизны. — 5. Мыльные пленки постоянной положительной кривизны и постоянной нулевой кривизны. — 6. Устойчивые и неустойчивые поверхности. — 7. Опыты Плато по устойчивости столба жидкости. — 8. Физическая реализация геликоида. — 9. Физическая реализация катеноида и его перестройки при изменении граничного контура. — 10. Изменение топологического типа минимальных поверхностей в зависимости от их устойчивости или неустойчивости. — 11. Двумерные минимальные поверхности в трехмерном пространстве и первый принцип Плато. — 12. Функционал площади, функционал Дирихле, гармонические отображения и конформные координаты. — 13. Особые, сингулярные точки минимальных поверхностей и три принципа Плато. — 14. Реализация минимальных поверхностей в живой природе.

1. О различных подходах к понятиям поверхности и границы поверхности. — 2. Гомологическая граница поверхности и роль группы коэффициентов. — 3. Удивительные примеры физических устойчивых минимальных поверхностей, ретрагирующихся тем не менее на свою границу. — 4. Когда мыльная пленка, затягивающая контур, не содержит в себе замкнутых мыльных пузырей? — 5. Минимальные конусы, связанные с особыми точками минимальных поверхностей. — 6. Многомерные минимальные конусы. — 7. Минимальные поверхности, инвариантные относительно действия групп Ли. — 8. Принцип Ферма, минимальные конусы и световые пучки. — 9. Проблема С. Н. Бернштейна. — 10. Комплексные подмногообразия как минимальные поверхности. — 11. Интегральные потоки. — 12. Бордизмы и классическая многомерная задача Плато. — 13. Существование минимума в каждом гомотопическом классе. — 14. Случай отсутствия решения задачи Дирихле для уравнения минимальной поверхности большой коразмерности. 15. Пример гладкой замкнутой незаузленной кривой в $\mathbb{R}^3$, которая ограничивает только минимальные поверхности большого рода.

1. Оценка снизу объемов минимальных поверхностей. — 2. Коэффициент деформации векторного поля. — 3. Замкнутые поверхности нетривиального топологического типа и наименьшего объема. — 4. Случай отсутствия локальных минимумов функционала Дирихле в классе отображений однородных пространств в произвольное риманово многообразие. — 5. Случай отсутствия локальных минимумов функционала Дирихле в классе отображений произвольного риманова многообразия в однородное пространство.



В рамках современной геометрии выделяется круг задач, расположенных на стыке нескольких научных областей: топологии, вариационного исчисления, алгебры, теории дифференциальных уравнений. В настоящей книге рассматриваются некоторые важные топологические вариационные задачи (как решенные, так и нерешенные), для исследования которых применяются современные аналитико-топологические методы.

В книге развиваются три взаимосвязанные темы. Первая из них носит скорее учебный характер (см. гл. 1) и посвящена изложению некоторых важнейших понятий топологии: гомологий и когомологий, теории препятствий к продолжению непрерывных отображений, расслоений и их простейших свойств. Этот раздел не носит самостоятельного характера, поскольку предназначен для обслуживания остальных глав книги, где перечисленные топологические понятия играют важную роль. В связи с этим изложение вспомогательного топологического материала максимально упрощено, громоздкие технические и формальные рассуждения заменены точными ссылками на соответствующую литературу. Это позволило сосредоточить основное внимание на выявлении наглядных геометрических механизмов, лежащих в основе исследуемых задач.

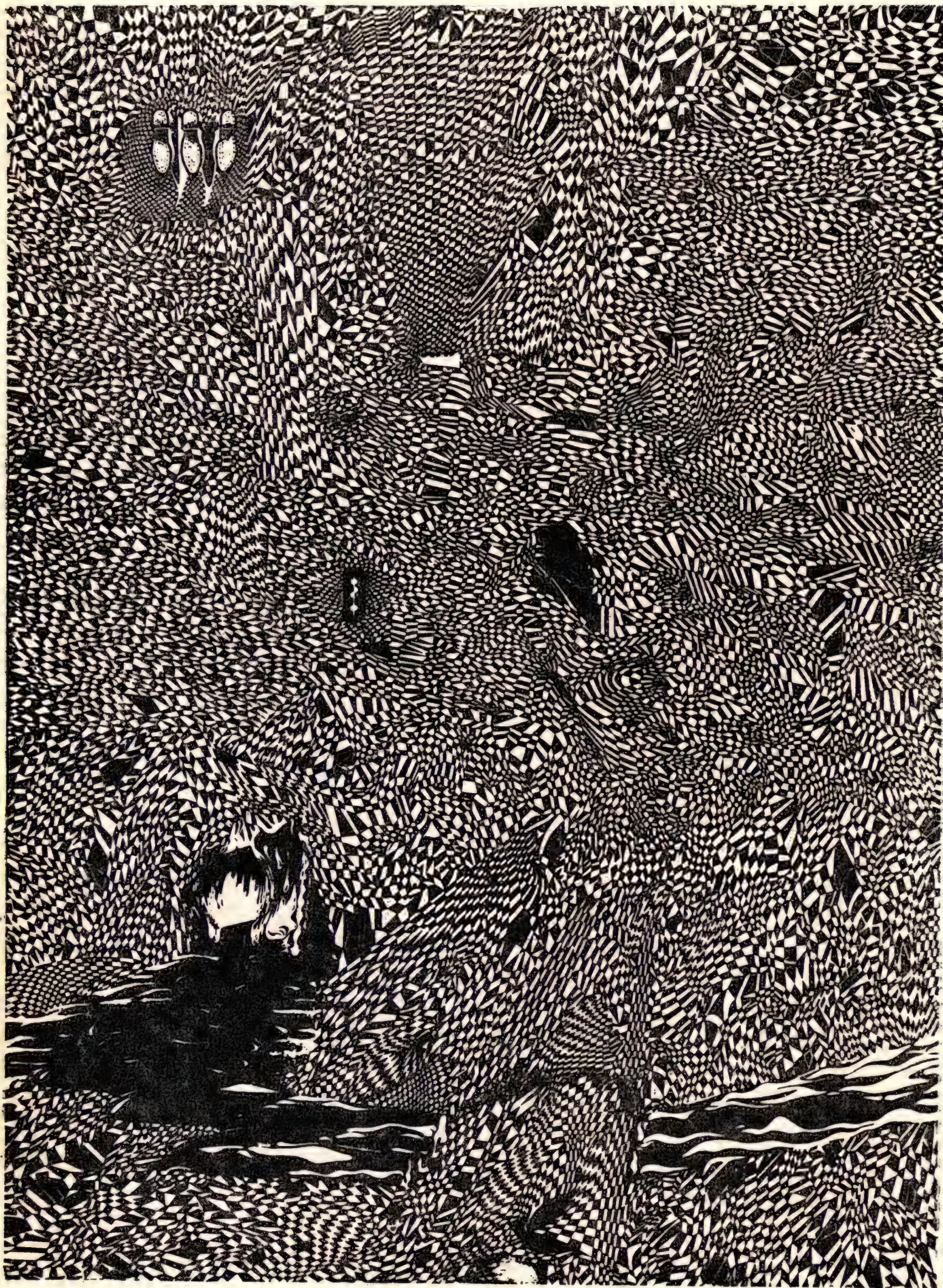
Второй круг вопросов, разрабатываемых в книге, объединен тем, что в них немаловажную (а иногда и основную) роль играют функции Морса на гладких многообразиях, т. е. функции, все критические точки которых невырождены. При этом мы сосредоточили основное внимание не на классической теории Морса, достаточно хорошо отраженной в учебной литературе, а на более современных аспектах топологии многообразий, связанных с функциями Морса. В частности, в гл. 2 затрагиваются вопросы теории Морса многозначных функций, полезной при

решении некоторых задач современной математической физики. Одной из важных топологических задач, в которых функции Морса играют существенную роль, является описание многообразий малых размерностей. В гл. 3 мы постарались, насколько это возможно в рамках краткого изложения, ввести читателя в круг вопросов, концентрирующихся около знаменитой проблемы Пуанкаре (гомеоморфно ли стандартной трехмерной сфере односвязное замкнутое трехмерное многообразие) и ее многомерных аналогов. В частности, мы излагаем новый и чрезвычайно простой алгоритм распознавания стандартной трехмерной сферы в классе диаграмм Хегора рода два, основанный на существовании так называемой «разбивающей вершины». В заключение мы приводим схему решения четырехмерной проблемы Пуанкаре, полученного Фридманом в 1981 г. Изложение пункта 2 в параграфе 1 и параграфа 3 в гл. 3 основано на материале, любезно предоставленном С. В. Матвеевым, за что автор выражает ему искреннюю благодарность.

Третья тема, изложению которой посвящена гл. 4, является, возможно, наиболее синтетической из всех других, затронутых в книге. Здесь речь идет о минимальных поверхностях, понимаемых в широком смысле. Наряду с классическими минимальными поверхностями, реализующими (локально) минимум площади, мы рассматриваем многомерные минимальные поверхности, экстремали функционала Дирихле, стратифицированные экстремали стратифицированных функционалов, интегральные потоки и т. д. Разнообразные топологические задачи, связанные с этими понятиями и их современными приложениями, объединены тем, что все они в той или иной форме используют топологические факты различного уровня сложности, изложенные в гл. 1. Отметим следующую характерную черту этого последнего раздела книги. Многие глубокие и по-настоящему нетривиальные геометрические утверждения из современного вариационного исчисления можно продемонстрировать на уровне реальных физических опытов, первые из которых были проведены еще в XIX в. и до сих пор сохраняют свое значение. В частности, мы рассказываем о некоторых опытах Плато с мыльными пленками. Здесь широко используются графические иллюстрации, позволяющие в компактной форме наглядно продемонстрировать нужные физические эффекты.

Книга рассчитана на широкие круги читателей, поэтому мы постарались свести к минимуму базисный набор фактов, необ-

ходимый для понимания материала, излагаемого в книге. Однако следует отметить, что предварительное знакомство с основными геометрическими понятиями, изложенными в стандартной учебной литературе, например в [1, 4, 5, 6], окажется, несомненно, полезным для более быстрого проникновения в существо некоторых из рассматриваемых вопросов.



§ 1. Группы сингулярных и клеточных гомологий

1. Сингулярные цепи и группы гомологий.

Как показали результаты, полученные в последнее время в связи с решением различных топологических вариационных задач, большую роль в них играют группы гомологий. Кратко напомним некоторые необходимые для нас сведения об этих группах и приведем также алгебраические и топологические конструкции, которые будут использоваться в дальнейшем.

Рассмотрим евклидово пространство $\mathbb{R}^{k+1}$, отнесенное к декартовым координатам $x_1, \dots, x_{k+1}$, и рассмотрим стандартный симплекс Δ^k размерности k , задаваемый так: $x_1 + \dots + x_{k+1} = 1$, $x_1 \geq 0, \dots, x_{k+1} \geq 0$.

Определение 1. Сингулярным симплексом f^k размерности k топологического пространства X называется непрерывное отображение f стандартного симплекса Δ^k в пространство X . Целочисленной k -мерной сингулярной цепью c пространства X называется формальная линейная комбинация сингулярных симплексов f^k пространства X с целыми коэффициентами, лишь конечное число которых отлично от нуля: $c = \sum_i a_i f_i^k$. Множество

всех k -мерных цепей, очевидно, превращается в абелеву группу по сложению $C_k(X)$, называемую группой k -мерных цепей пространства X . Граничным оператором $\partial_k : C_k(X) \rightarrow C_{k-1}(X)$ называется гомоморфизм, определяемый на образующих следующей формулой: $\partial_k f^k = \sum_{i=0}^k (-1)^i f_i^{k-1}$, где f_i^{k-1} —

ограничение отображения f^k на i -ю грань симплекса Δ^k .

Ясно, что $\partial_k \partial_{k+1} \equiv 0$, поэтому $\text{Ker } \partial_k \supset \text{Im } \partial_{k+1}$.

Определение 2. Группой k -мерных сингулярных гомологий $H_k(X)$ пространства X называется фактор-группа $\text{Ker } \partial_k / \text{Im } \partial_{k+1}$. Элементы подгруппы $\text{Ker } \partial_k$ называются циклами, а элементы подгруппы $\text{Im } \partial_{k+1}$ — границами.

Из определения групп сингулярных гомологий легко следует важное свойство, заключающееся в том, что эти группы топологически и гомотопически инвариантны, т. е. не меняются при гомеоморфизме пространства или при замене его на гомотопически эквивалентное пространство. Совокупность групп $C_k(X)$ и связывающих их гомоморфизмов ∂_k естественно организуется в последовательность:

$$\dots \rightarrow C_k \xrightarrow{\partial_k} C_{k-1} \rightarrow \dots \rightarrow C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\varepsilon} \mathbb{Z} \rightarrow 0,$$

где $\partial_k \partial_{k+1} \equiv 0$ и ε — эпиморфизм. Эта последовательность называется цепным комплексом. Отображение одного цепного комплекса в другой, коммутирующее с гомоморфизмами ∂_k , называется цепным отображением.

Будем говорить, что задана цепная гомотопия D комплекса C в комплекс C' , связывающая два отображения φ и ψ комплекса C (где $\varphi = \{\varphi_k : C_k \rightarrow C'_k\}$, $\psi = \{\psi_k : C_k \rightarrow C'_k\}$), если задана совокупность гомоморфизмов $D_k : C_k \rightarrow C'_{k+1}$ таких, что для каждого k имеем: $D_{k-1} \partial_k + \partial'_{k+1} D_k = \varphi_k - \psi_k$. Два цепных отображения φ и ψ , связанные цепной гомотопией, называются иногда цепно-гомотопными. Из определения групп гомологий следует, что гомотопные цепные отображения индуцируют одинаковые отображения групп гомологий $H_k(C)$ в $H_k(C')$, где $H_k(C) = \text{Ker } \partial_k / \text{Im } \partial_{k+1}$.

Для вычисления гомологий чрезвычайно полезна так называемая точная последовательность пары. Пусть даны топологические пространства X, Y , где Y — замкнутое подпространство в X . Ясно, что $C_k(Y) \subset C_k(X)$, и можно рассмотреть группу относительных цепей $C_k(X, Y) = C_k(X) / C_k(Y)$. Так как граничный оператор действует следующим образом: $\partial_k : C_k(X) \rightarrow C_{k-1}(X)$, $\partial_k : C_k(Y) \rightarrow C_{k-1}(Y)$, то он индуцирует некоторый оператор $C_k(X, Y) \rightarrow C_{k-1}(X, Y)$, который мы для простоты обозначим тем же символом ∂_k . Можно определить группы $\text{Ker } \partial_k = Z_k(X, Y) \supset \supset B_k(X, Y) = \text{Im } \partial_k$ (относительные циклы и относительные границы), что позволяет рассмотреть фактор-группу $H_k(X, Y) = Z_k(X, Y) / B_k(X, Y)$, называемую группой относительных k -мерных гомологий пространства X по модулю подпространства Y . Легко проверяется, что для групп относительных гомологий также имеют место топологическая и гомотопическая инвариантность. Перейдем теперь к построению нового оператора $\partial : H_k(X, Y) \rightarrow H_{k-1}(Y)$. Пусть $z_k \in C_k(X, Y)$ — относительный цикл и пусть $\tilde{z}_k \in C_k(X)$ — произвольный его представитель в классе смежности. Так как $\partial_k z_k = 0$, то $\partial_k \tilde{z}_k \in C_{k-1}(Y)$. Этот абсолютный цикл мы обозначим через $\partial z_k \in Z_{k-1}(Y)$. Класс гомологий этого цикла не зависит от выбора представителя из класса гомологий цикла z_k . Это и определяет нам некоторый гомоморфизм (оператор) $\partial : H_k(X, Y) \rightarrow H_{k-1}(Y)$, который мы также будем обозначать ∂ и называть граничным (рис. 1). Обозначим, далее, через $i : Y \rightarrow X$ вложение, тогда оно индуцирует гомоморфизм $i_* : H_k(Y) \rightarrow H_k(X)$. Поскольку любой абсолютный цикл можно рассматривать как относительный (по модулю подпространства Y), то возникает еще одно естественное отображение $j_* : H_k(X) \rightarrow H_k(X, Y)$.

Теорема 1. Следующая последовательность групп и гомоморфизмов является точной, т. е. в каждом ее члене образ входящего гомоморфизма совпадает с ядром исходящего гомоморфизма:

$$\dots \rightarrow H_{k+1}(X, Y) \xrightarrow{\partial} H_k(Y) \xrightarrow{i_*} H_k(X) \xrightarrow{j_*} H_k(X, Y) \xrightarrow{\partial} H_{k-1}(Y) \rightarrow \dots$$

Доказательство. Проверим, например, точность в члене $H_k(X, Y)$. Если $\partial\alpha=0$, $\alpha\in H_k(X, Y)$, то ∂z_k гомологично нулю в Y (где z_k — представитель α , т. е. $z_k\in Z_k(X, Y)$). Но тогда, добавляя к z_k k -мерную цепь, осуществляющую эту гомологию в подпространстве Y , мы получаем k -мерную цепь, являющуюся циклом уже с точки зрения объемлющего пространства X , т. е. мы представили исходный элемент α в виде образа некоторого элемента β , т. е. $\alpha=j_*\beta$, $\beta\in H_k(X)$ (рис. 2). Итак, $\text{Ker } \partial\subset \text{Im } j_*$. Обратное включение $\text{Ker } \partial\supset \text{Im } j_*$ следует

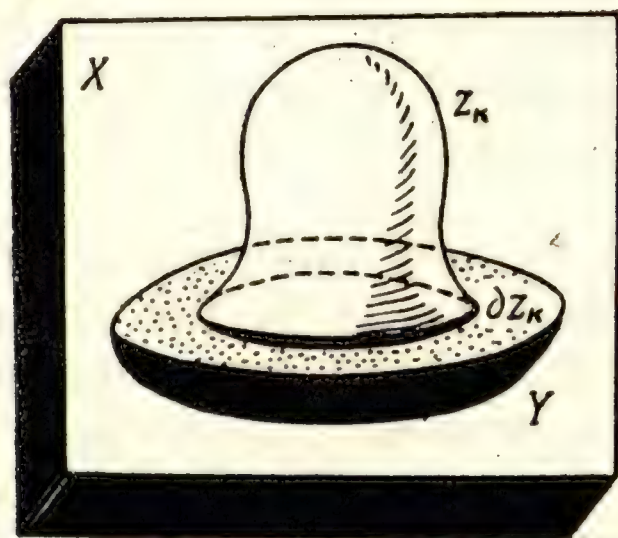


Рис. 1

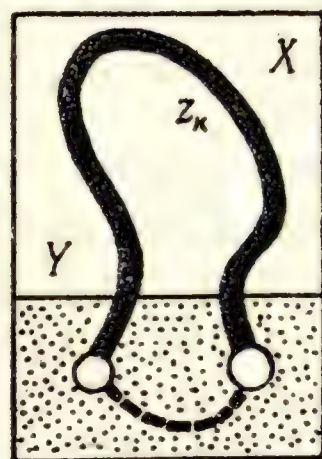


Рис. 2

из того, что любой абсолютный цикл в X , рассматриваемый затем как относительный, имеет нулевую границу в Y . Точность последовательности в остальных членах проверяется аналогичными рассуждениями.

Для дальнейшего полезно знать следующие простые свойства точных последовательностей (доказательства мы оставляем читателю в качестве упражнения):

1) Последовательность $0\rightarrow A\rightarrow 0$ точна тогда и только тогда, когда $A=0$. 2) Последовательность $0\rightarrow A\overset{\alpha}{\rightarrow} B\rightarrow 0$ точна тогда и только тогда, когда группы A и B изоморфны и гомоморфизм α является изоморфизмом. 3) Последовательность $0\rightarrow A\overset{i}{\rightarrow} B\overset{\pi}{\rightarrow} C\rightarrow 0$ точна тогда и только тогда, когда группа A является подгруппой группы B , причем гомоморфизм $i: A\rightarrow B$ — включение (мономорфизм), $C=B/A$ (фактор-группа) и $\pi: B\rightarrow B/A=C$ — естественная проекция на фактор-группу.

Оказывается, существует редукция относительных гомологий к абсолютным.

2. Клеточные комплексы, барицентрические подразделения.

Если X, Y линейно связны, то $H_0(X, Y)=0$. Топологическое пространство X называется клеточным комплексом, если оно представляется в виде объединения непересекающихся подмножеств σ^k , называемых клетками, причем замыкание каждой клетки σ^k является образом замкнутого k -мерного диска D^k при некотором непрерывном отображении (называемом характеристическим), являющемся гомеоморфизмом на внутренности

диска. Далее, граница каждой клетки (т. е. образ границы диска при характеристическом отображении) должна содержаться в объединении конечного числа клеток меньших размерностей. Наконец, требуется, чтобы подмножество в X было замкнуто тогда и только тогда, когда замкнуты все полные прообразы (при характеристических отображениях) пересечений этого подмножества со всеми клетками.

Комплекс называется конечным, если он состоит из конечного числа клеток. Объединение клеток размерностей, не превосходящих k , называется k -мерным остовом комплекса. Будем говорить, что пара пространств (X, Y) является клеточной парой, если X и Y — клеточные комплексы и замкнутое подпространство Y — клеточный подкомплекс в X . Ограничимся рассмотрением конечных комплексов. Через X/Y обозначим пространство, получающееся из X отождествлением замкнутого подпространства Y в одну точку.

Теорема 2. Пусть (X, Y) — клеточная пара. Тогда $H_k(X, Y) = H_k(X/Y)$ при $k \neq 0$.

Доказательство. Обозначим через CY конус над Y , т. е. пространство, получающееся из цилиндра $Y \times I$ стягива-

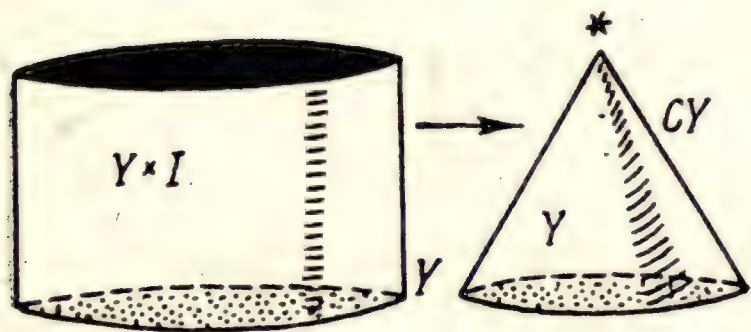


Рис. 3

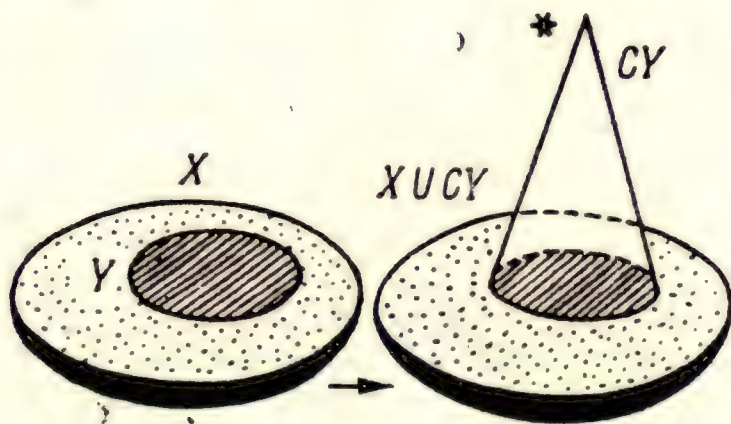


Рис. 4

нием в одну точку верхнего основания (рис. 3). Построим новое пространство $X \cup CY$, т. е. отождествим подпространство Y в X с основанием Y в конусе CY (рис. 4). Так как X, Y — конечные комплексы, то, стягивая конус CY по себе в точку, мы получаем гомотопическую эквивалентность: $X \cup CY \approx X/Y$ (рис. 5). Таким образом, для доказательства теоремы достаточно установить, что при $k > 0$ выполняются соотношения: $H_k(X, Y) = H_k(X \cup CY)$. Из точной последовательности пары $(X \cup CY, *)$, где $*$ — точка, вершина конуса, следует, что $H_k(X \cup CY, *) = H_k(X \cup CY)$ при $k > 0$, поэтому мы должны доказать, что $H_k(X, Y) = H_k(X \cup CY, *)$, при $k > 0$. Прежде чем переходить к доказательству, нам потребуется одно техническое рассуждение, связанное с так называемыми барицентрическими подразделениями.

Пусть Δ^k — стандартный симплекс. Его барицентрическое подразделение $\beta\Delta^k$ определим по индукции так. Если $k=1$, то $\beta\Delta^1$ получается добавлением новой вершины в середине отрезка Δ^1 . Если $k=2$, то $\beta\Delta^2$ получается путем соединения центра

треугольника с его вершинами и с серединами его сторон (рис. 6). И наконец, при произвольном k подразделение $\beta\Delta^k$ получается так: отмечаем в Δ^k его центр и разбиваем симплекс Δ^k на пирамиды с вершиной в этом центре и с основаниями —

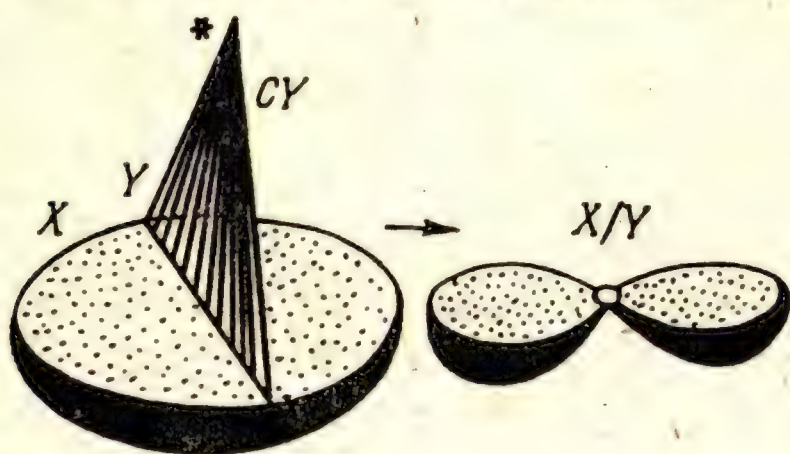


Рис. 5

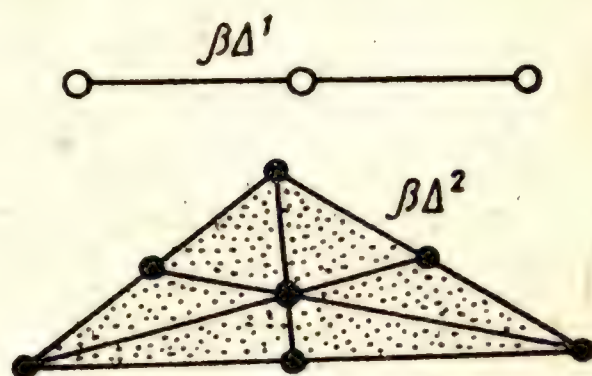


Рис. 6

симплексами барицентрического подразделения границы симплекса Δ^k . Пусть теперь $f: \Delta^k \rightarrow X$ — произвольный сингулярный симплекс пространства X . Через βf мы обозначим цепь, равную сумме всех сингулярных симплексов, получающихся при ограничении исходного отображения f на k -мерные симплексы барицентрического подразделения

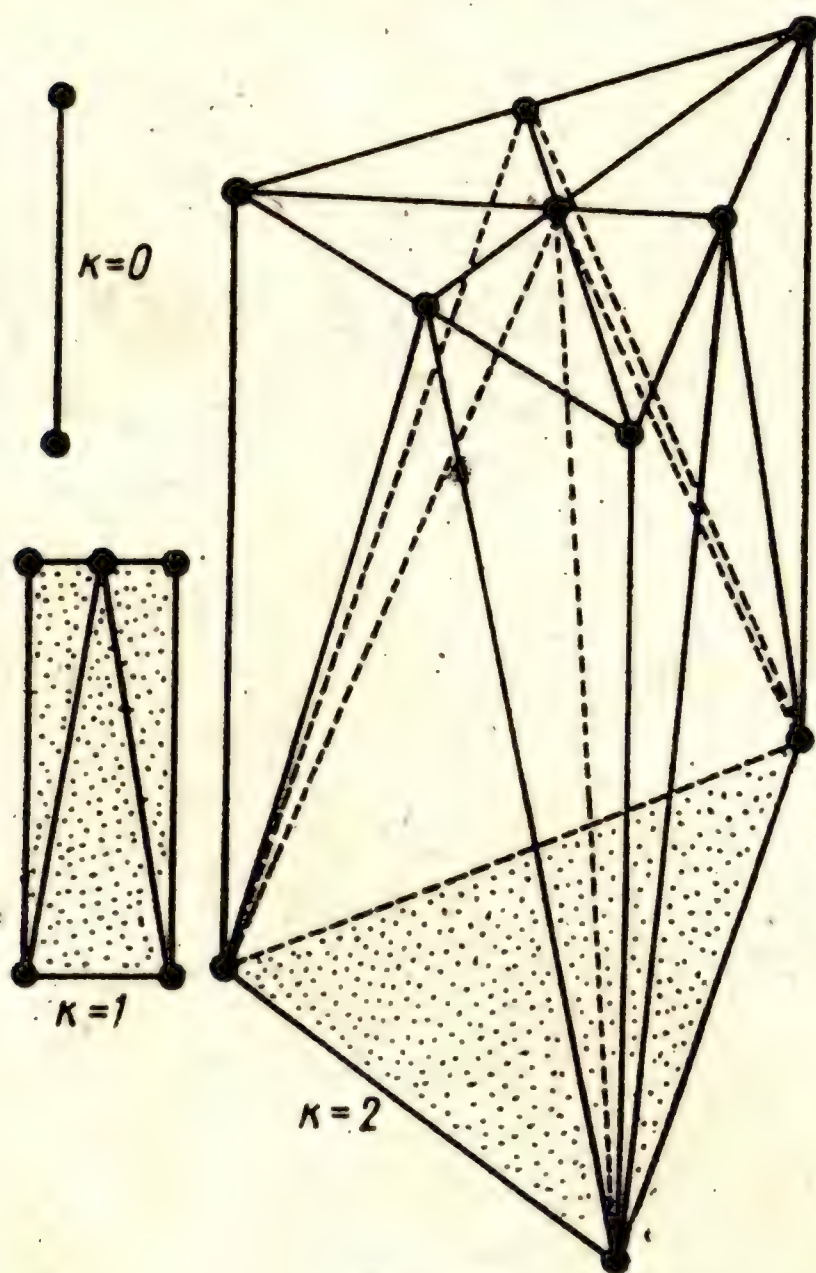


Рис. 7

барицентрического подразделения $\beta\Delta^k$ исходного симплекса Δ^k . Сопоставляя каждому сингулярному симплексу f (т. е. элементарной цепи в $C_k(X)$) сингулярную цепь βf , мы получаем гомоморфизм $\beta_k: C_k(X) \rightarrow C_k(X)$. Мы утверждаем, что совокупность отображений $\beta = \{\beta_k\}$ определяет цепное отображение β цепного комплекса $C(X)$ в себя, причем оно цепно-гомотопно тождественному. То, что отображение β является цепным, следует из того, что ограничение формальной суммы сингулярных симплексов на барицентрическое подразделение равно формальной сумме этих ограничений. Построим в явном виде цепную гомотопию D , связывающую отображение β с тождественным. Определим при каждом $k > 0$ разбиение

призмы $\Delta^k \times I$ в сумму симплексов следующим образом. На рис. 7 это разбиение показано при $k=0, 1, 2$. Индуктивный процесс разбиения $\Delta^k \times I$ можно описать так. Если разбиение уже определено при $q < k$, то на части границы $\Delta^k \times I$, а именно

на объединении (на «стакане») $(\Delta^k \times 0) \cup (\partial \Delta^k \times I)$ разбиение нужно взять так, чтобы $\Delta^k \times 0$ был стандартным симплексом, а $\partial \Delta^k \times I$ разбит в соответствии с индуктивным процессом при $q < k$. После этого разбиваем всю призму на k -мерные симплексы, основаниями которых являются $(k-1)$ -мерные симплексы указанного разбиения части границы, а вершиной — центр верхней грани. Ясно, что при этом верхняя грань подвергнется барицентрическому подразделению.

Пусть теперь $f: \Delta^k \rightarrow X$ — сингулярный симплекс. Через $D_k f \in C_{k+1}(X)$ мы обозначим $(k+1)$ -мерную цепь, являющуюся суммой всех $(k+1)$ -мерных сингулярных симплексов, получающихся при ограничении отображения $\Phi: \Delta^k \times I \rightarrow X$, где $\Phi(x, t) = f(x)$, на симплексы построенного нами выше разбиения ци-

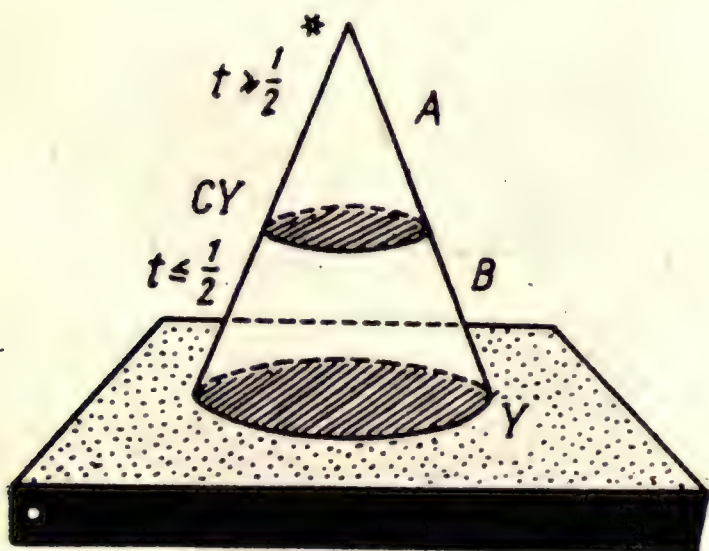


Рис. 8

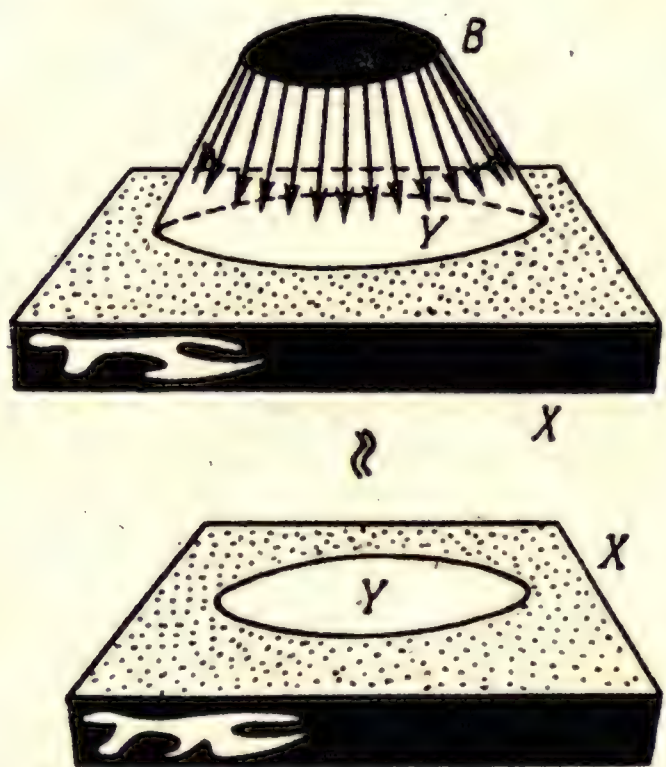


Рис. 9

линдра $\Delta^k \times I$. Мы получаем гомоморфизмы $D_k: C_k(X) \rightarrow C_{k+1}(X)$, которые и определяют искомую цепную гомотопию, соединяющую тождественное отображение комплекса $C(X)$ на себя с барицентрическим отображением β . Доказательство того, что отображение $\{D_k\}$ задают цепную гомотопию, мы предоставляем читателю. Переходим теперь к доказательству теоремы. Рассмотрим отображение вложения пар $(X, Y) \rightarrow (X \cup CY, CY)$. Оно индуцирует гомоморфизм $\alpha: H_k(X, Y) \rightarrow H_k(X \cup CY, CY) = H_k(X \cup CY, *)$. Мы воспользовались тем, что $CY \approx *$, т. е. конус стягивается в точку. Докажем, что α — эпиморфизм. Пусть $z \in Z_k(X \cup CY, CY)$ — некоторый цикл. Мы должны найти в группе $Z_k(X, Y)$ цикл, который при указанном отображении перейдет в цикл, гомологичный z в $Z_k(X \cup CY, CY)$. Представим конус CY в виде объединения двух подмножеств: конуса A — части CY , состоящей из точек, для которых $t \geq 1/2$, и усеченного конуса B — части CY , состоящей из точек, для которых $t \leq 1/2$ (рис. 8). Измельчая путем барицентрических подразделений сингулярные симплексы, составляющие цикл z , мы получаем

каждый раз гомологичные ему циклы, но состоящие из все более и более мелких симплексов. В силу конечности числа симплексов и компактности X, Y существует такое достаточно мелкое барицентрическое подразделение, что если какой-либо сингулярный симплекс этого подразделения пересекается с A , то он полностью лежит в конусе CY . Здесь мы рассматриваем образ стандартного симплекса при отображении, задающем сингулярный симплекс. Этот цикл z' гомологичен z . Выбросим из цикла z' все те симплексы, которые пересекаются с A . Поскольку эта операция производится в пределах конуса CY , то с точки зрения относительных гомологий новый цикл z'' остается в том же классе гомологий, что и z' , а следовательно, и z . Итак, $z'' \in H_k(X \cup B, B)$. В то же время в силу гомотопической инвариантности групп гомологий имеем: $H_k(X \cup B, B) = H_k(X, Y)$, поскольку $(X \cup B, B) \approx (X, Y)$ (рис. 9). Итак, мы в явном виде предъявили некоторый цикл $z'' \in H_k(X, Y)$, который

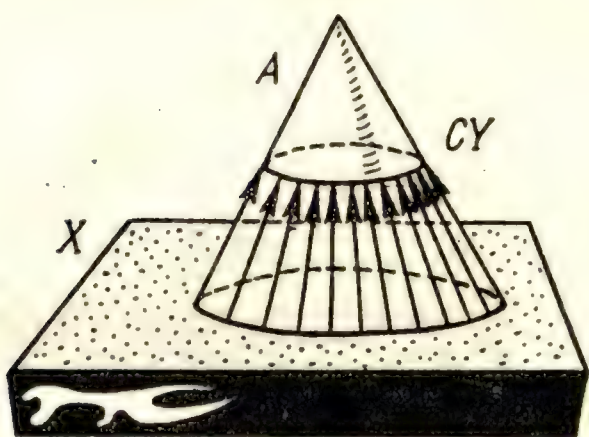


Рис. 10

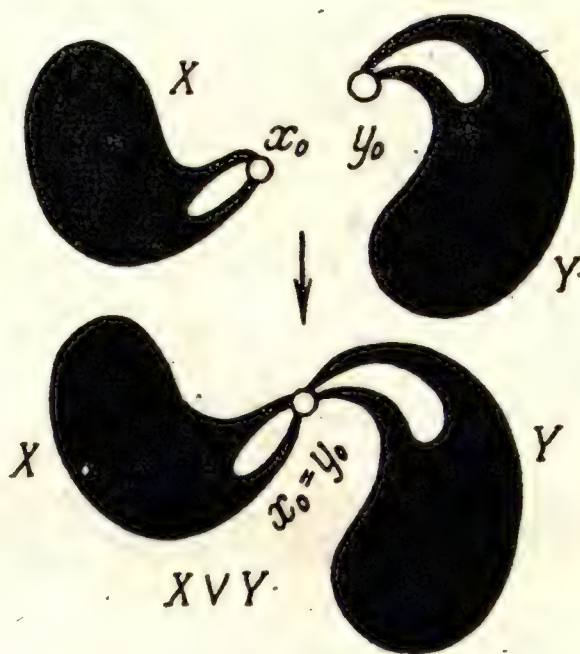


Рис. 11

при гомоморфизме α переходит в цикл $z \in H_k(X \cup CY, CY)$. Тем самым доказана эпиморфность α . Доказательство мономорфности проводится по аналогичной схеме и предоставляется читателю; см. рис. 10.

3. Клеточные гомологии и вычисление сингулярных гомологий сферы.

Выше мы описали один из способов введения групп гомологий, используя сингулярные симплексы. Однако для клеточных комплексов можно определить так называемые клеточные гомологии, совпадающие, как оказывается, с сингулярными, но обладающие одним важным преимуществом: клеточные гомологии существенно легче вычислять. Даже в простейшем случае, когда пространство состоит из одной точки, вычисление сингулярных гомологий требует некоторых (впрочем элементарных) рассуждений. Если же пространство устроено более сложно, то подсчет его сингулярных гомологий резко усложняется. На практике особенно часто используются поэтому

именно клеточные гомологии. Мы введем их, опираясь на уже знакомое нам понятие сингулярных гомологий, что позволит затем не только доказать совпадение групп сингулярных и клеточных гомологий, но и предъявить наглядную схему их вычисления.

Подсчитаем сингулярные гомологии сферы. Поскольку нульмерная сфера S^0 состоит из пары точек, то мы получаем: $H_k(S^0) = 0$ при $k \neq 0$ и $H_0(S^0) = \mathbb{Z} \oplus \mathbb{Z}$.

Лемма 1. Группы сингулярных гомологий n -мерной сферы S^n , где $n > 0$, имеют вид: $H_k(S^n) = 0$ при $k \neq 0, n$ и $H_k(S^n) = \mathbb{Z}$ при $k = 0, n$.

Доказательство вытекает из рассмотрения точной гомологической последовательности пары (D^{n+1}, S^n) и из того факта, что диск стягивается в точку.

Напомним определение букета двух топологических пространств. Пусть в пространствах X и Y выделены точки x_0 и y_0 соответственно. Построим новое пространство $X \vee Y$, отождествив эти точки (рис. 11). Никакие другие точки не отождествляются.

Лемма 2. Пусть $X = \bigvee S_i^n$ — букет n -мерных сфер S_i^n , занумерованных индексом i , где $1 \leq i \leq N$. Если $n > 0$, $k > 0$, то имеет место изоморфизм $H_k(X) = \bigoplus_i H_k(S_i^n) = \mathbb{Z} \oplus \dots \oplus \mathbb{Z}$ (N раз).

Для доказательства достаточно рассмотреть точную последовательность пары $(\bigvee D_i^n, \bigvee \partial D_i^n)$, причем

$$\bigvee D_i^n / \bigvee \partial D_i^n = \bigvee S_i^n$$

(рис. 12).

Отметим полезный для дальнейшего факт. В качестве образующей в группе $\mathbb{Z} = H_n(D^n, S^{n-1})$ можно взять класс гомологий простейшей сингулярной цепи $1 \cdot f$, где $f: \Delta^n \rightarrow D^n$ — гомеоморфизм симплекса на диск. Следовательно, ориентацию сферы можно задавать, фиксируя образующую в группе $\mathbb{Z} = H_n(S^n)$. Изменение ориентации эквивалентно замене элемента 1 на элемент -1 .

Определим теперь группы клеточных цепей. Пусть X — конечный клеточный комплекс. Попробуем вычислить его сингулярные гомологии в терминах клеток и их характеристических отображений, т. е. в тех терминах, в которых нам задан комплекс. Множество всех k -мерных клеток комплекса X обозначим через X_k . Пусть X^k — k -мерный остов комплекса X . Будем считать, что фиксированы ориентации всех клеток. Занумеруем все k -мерные клетки и пусть A_k — множество индексов. Тогда

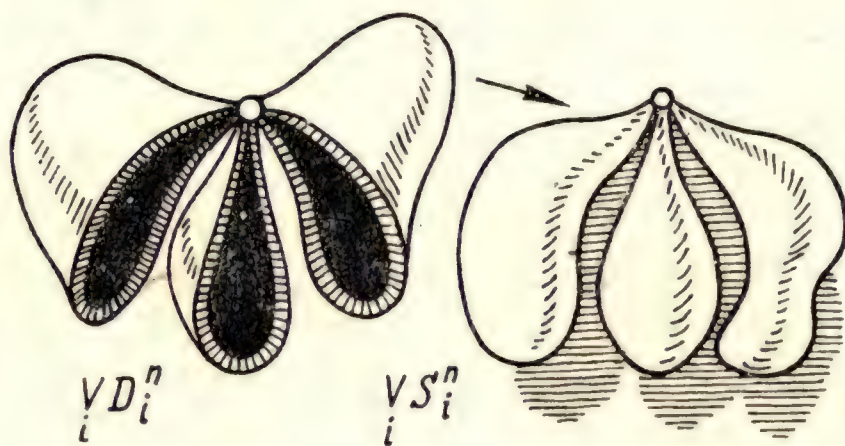


Рис. 12

в силу леммы 2 $H_i(X^k, X^{k-1}) = H_i(\bigvee_{\alpha \in A_k} S_\alpha^k) = \begin{cases} 0 & \text{при } i \neq k, \\ P_k(X) & \text{при } i = k. \end{cases}$

Здесь через $P_k(X)$ обозначена свободная абелева группа, образующие которой находятся во взаимно-однозначном соответствии с A_k . Поскольку элементы этой группы естественно отождествляются с линейными комбинациями вида $\sum_{\alpha} a_{\alpha} \sigma_{\alpha}^k$, где

σ_{α}^k — k -мерные клетки комплекса X , то группа $P_k(X)$ — конечно-порожденная. Эту группу назовем группой клеточных k -мерных цепей комплекса X . Группы $P_k(X)$ и $C_k(X)$ неизоморфны в общем случае. Прежде чем двигаться дальше, рассмотрим так называемую точную гомологическую последовательность тройки — вариант последовательности пары. Пусть (X, Y, Z) — тройка пространств, где Y и Z замкнуты в X . Рассмотрим два вложения: $(Y, Z) \rightarrow (X, Z)$ и $(X, Z) \rightarrow (X, Y)$; пусть $\partial: H_k(X, Y) \rightarrow H_{k-1}(Y, Z)$ — граничный гомоморфизм, порожденный гомоморфизмом $\partial: H_k(X, Y) \rightarrow H_{k-1}(Y)$, определение которого см. выше, и тем обстоятельством, что каждый абсолютный цикл из $H_{k-1}(Y)$ можно рассматривать как относительный по $\text{mod } Z$, т. е. как элемент группы $H_{k-1}(Y, Z)$. Тогда получаем последовательность:

$$\dots \xrightarrow{\partial} H_k(X, Y) \xrightarrow{\partial} H_{k-1}(Y, Z) \xrightarrow{\partial} H_{k-1}(X, Z) \xrightarrow{\partial} H_{k-1}(X, Y) \xrightarrow{\partial} \dots$$

Проверка ее точности предоставляется читателю.

Возвращаясь к группам $P_k(X) = H_k(X^k/X^{k-1})$ и $P_{k-1}(X) = H_{k-1}(X^{k-1}/X^{k-2})$, мы можем рассмотреть точную последовательность тройки (X^k, X^{k-1}, X^{k-2}) . В данный момент из нее нам нужен только гомоморфизм $H_k(X^k, X^{k-1}) \rightarrow H_{k-1}(X^{k-1}, X^{k-2})$, который в наших новых обозначениях записывается так: $P_k(X) \rightarrow P_{k-1}(X)$. Обозначая его через ∂_k , мы получаем цепной

комплекс $\{P_k(X), \partial_k\}$, а именно $\dots \xrightarrow{\partial_k} P_k(X) \xrightarrow{\partial_k} P_{k-1}(X) \xrightarrow{\partial_k} \dots$.

Как и для любого цепного комплекса, определены его группы гомологий, т. е. группы $\text{Ker } \partial / \text{Im } \partial$, называемые клеточными гомологиями комплекса. Оказывается, существует канонический изоморфизм между гомологиями описанного цепного комплекса и сингулярными гомологиями X . Это и есть основное утверждение настоящего параграфа, позволяющее сводить вычисление сингулярных гомологий к подсчету гомологий значительно более простого цепного комплекса. Оказывается, эта редукция настолько эффективна, что подавляющее большинство конкретных вычислений гомологий основано именно на этой теореме. В частности, из нее сразу следует, что группы клеточных гомологий гомотопически инвариантны и что группы сингулярных гомологий конечного комплекса конечно-порождены.

4. Теорема о совпадении сингулярных и клеточных гомологий конечного клеточного комплекса.

Теорема 3. Для конечного клеточного комплекса X группы сингулярных гомологий $H_k(X)$ и группы гомологий цепного

комплекса $\{P_k(X), \partial_k\}$, т. е. группы $\text{Ker } \partial_k / \text{Im } \partial_{k+1}$ (группы клеточных гомологий) изоморфны.

Сначала докажем некоторые вспомогательные утверждения.

Лемма 3. При $k > 1$ имеет место изоморфизм $H_k(X^{k+1}, X^{k-2}) = H_k(X)$.

Доказательство. Рассмотрим тройку комплексов $(X^{k+1}, X^{k-2}, X^{k-3})$ и соответствующую ей точную последовательность:

$$H_k(X^{k-2}, X^{k-3}) \rightarrow H_k(X^{k+1}, X^{k-3}) \rightarrow H_k(X^{k+1}, X^{k-2}) \rightarrow \\ \rightarrow H_{k-1}(X^{k-2}, X^{k-3}).$$

Крайние ее члены равны нулю, т. е. $H_k(X^{k+1}, X^{k-3}) = H_k(X^{k+1}, X^{k-2})$. Повторяя рассуждение для тройки $(X^{k+1}, X^{k-3}, X^{k-4})$, получаем, что $H_k(X^{k+1}, X^{k-3}) = H_k(X^{k+1}, X^{k-4})$. Двигаясь дальше по размерности, получаем следующую цепочку изоморфизмов:

$$H_k(X^{k+1}, X^{k-2}) = H_k(X^{k+1}, X^{k-3}) = H_k(X^{k+1}, X^{k-4}) = \dots \\ \dots = H_k(X^{k+1}, X^0) = H_k(X^{k+1})$$

при $k > 1$. Если остов X^0 состоит только из одной точки, то равенство верно и при $k = 1$. Любой конечный клеточный связный комплекс гомотопически эквивалентен такому конечному комплексу, нульмерный остов которого состоит ровно из одной точки. Для этого достаточно рассмотреть все нульмерные клетки исходного комплекса, соединить каждую из них непрерывным путем с одной выделенной вершиной (нульмерной клеткой *), причем эти пути должны лежать в одномерном остове X^1 . Затем осуществим гомотопию, показанную на рис. 13 и стягивающую все нульмерные клетки в одну.

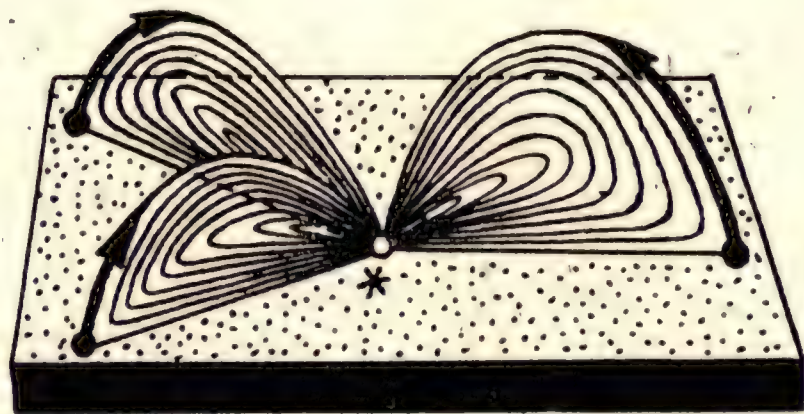


Рис. 13

Лемма 4. Имеет место изоморфизм: $H_k(X) = H_k(X^{k+1})$.

Доказательство. Докажем, что для любого $i < k+1$ имеет место равенство $H_i(X^{k+1}) = H_i(X^{k+2})$. В самом деле, рассмотрим точную последовательность пары (X^{k+2}, X^{k+1}) :

$0 = H_{i+1}(X^{k+2}, X^{k+1}) \rightarrow H_i(X^{k+1}) \rightarrow H_i(X^{k+2}) \rightarrow H_i(X^{k+2}, X^{k+1}) = 0$. Отсюда и следует требуемое равенство, в частности $H_k(X^{k+1}) = H_k(X^{k+2})$. Возвращаясь к доказательству леммы 3, мы получаем

$$H_k(X^{k+1}, X^{k-2}) = H_k(X^{k+1}) = H_k(X^{k+2}) = \dots = H_k(X),$$

что и требовалось.

Лемма 5. Имеет место изоморфизм: $\text{Ker } \partial_k / \text{Im } \partial_{k+1} =$

$= H_k(X^{k+1}, X^{k-2})$, где ∂_k — гомоморфизмы, определяющие комплекс групп клеточных цепей.

Доказательство. Рассмотрим коммутативную диаграмму:

$$\begin{array}{ccccccc}
 & & & & H_k(X^{k-1}, X^{k-2}) = 0 & & \\
 & & & & \downarrow \partial & & \\
 P_{k+1}(X) = H_{k+1}(X^{k+1}, X^k) & \xrightarrow{\partial} & H_k(X^k, X^{k-2}) & \xrightarrow{i_*} & H_k(X^{k+1}, X^{k-2}) & \rightarrow & H_k(X^{k+1}, X^k) = 0 \\
 \downarrow & & \downarrow j_* & & & & \\
 & \xrightarrow{\partial_{k+1}} & H_k(X^k, X^{k-1}) = P_k(X) & & & & \\
 & & \downarrow \partial_k & & & & \\
 & & H_{k-1}(X^{k-1}, X^{k-2}) = P_{k-1}(X) & & & &
 \end{array}$$

Здесь строка является отрезком точной последовательности тройки (X^{k+1}, X^k, X^{k-2}) , столбец — отрезок последовательности тройки (X^k, X^{k-1}, X^{k-2}) , а гомоморфизмы i_* , j_* индуцированы соответствующими вложениями пар i , j . Коммутативность диаграммы означает, что $\partial_{k+1} = j_* \partial$. Напомним, что $P_\alpha(X) = H_\alpha(X^\alpha, X^{\alpha-1})$. Поскольку строка и столбец — фрагменты точных последовательностей, то i_* — эпиморфизм, а j_* — мономорфизм. Отсюда

$$H_k(X^{k+1}, X^{k-2}) = H_k(X^k, X^{k-2}) / \text{Ker } i_* = H_k(X^k, X^{k-2}) / \text{Im } \partial.$$

Поскольку j_* — мономорфизм, то

$$\begin{aligned}
 H_k(X^k, X^{k-2}) / \text{Im } \partial &= j_* H_k(X^k, X^{k-2}) / j_* \text{Im } \partial = \\
 &= \text{Im } j_* / \text{Im } j_* \partial = \text{Ker } \partial_k / \text{Im } j_* \partial = \text{Ker } \partial_k / \text{Im } \partial_{k+1}.
 \end{aligned}$$

Здесь мы использовали тождества: $\text{Im } j_* = \text{Ker } \partial_k$ (в силу точности), $j_* \partial = \partial_{k+1}$ (коммутативность диаграммы). Итак,

$$H_k(X^{k+1}, X^{k-2}) = \text{Ker } \partial_k / \text{Im } \partial_{k+1}.$$

Лемма доказана.

Доказательство теоремы 3. Из лемм 3 и 5 получаем: $H_k(X) = \text{Ker } \partial_k / \text{Im } \partial_{k+1} = H_k(X^{k+1}, X^{k-2})$, что и требовалось.

5. Геометрическое определение групп клеточных гомологий.

Нам нужно прояснить геометрический смысл оператора ∂_k в цепном комплексе $\{P_k(X), \partial_k\}$. Рассмотрим две клетки σ^{k-1} и σ^k в X . Будем считать, что их ориентации фиксированы и что с этими ориентациями согласованы характеристические отображения $\chi: D^k \rightarrow X$, $\bar{\chi}: D^{k-1} \rightarrow X$. Рассмотрим непрерывное отображение $\partial D^k = S^{k-1} \xrightarrow{\chi} X^{k-1}/X^{k-2}$, т. е. отобразим границу клетки в фактор-пространство $(k-1)$ -мерного остова по $(k-2)$ -мерному остову. Это фактор-пространство гомеоморфно букету $(k-1)$ -мерных сфер. Поскольку в $(k-1)$ -мерном остове X^{k-1} нами выделена клетка σ^{k-1} , то она при указанной факторизации $X^{k-1} \xrightarrow{\alpha} X^{k-1}/X^{k-2}$ отобразится на некоторую сферу S^{k-1} из букета X^{k-1}/X^{k-2} . Итак, в букете сфер выделилась одна сфера, что позволяет определить естественное проектирование всего букета на эту сферу.

А именно: выделенная сфера остается на месте, а все остальные отображаются в отмеченную точку (рис. 14). Построенное отображение границы шара D^k в букет сфер с последующим проектированием на выделенную сферу определяет непрерывное отображение $S^{k-1} \rightarrow S^{k-1}$, в частности задает гомоморфизм группы $Z = H_{k-1}(S^{k-1})$ в себя. Каждый гомоморфизм $Z \rightarrow Z$ однозначно определяется целым числом m — образом единицы группы Z . Это число называется степенью отображения. В том случае, когда построенное нами отображение является гладким (так будет во многих примерах, рассматриваемых ниже), число m совпадает с обычной степенью гладкого отображения, определяемого для отображений ориентируемых замкнутых многообразий одинаковой размерности [1].

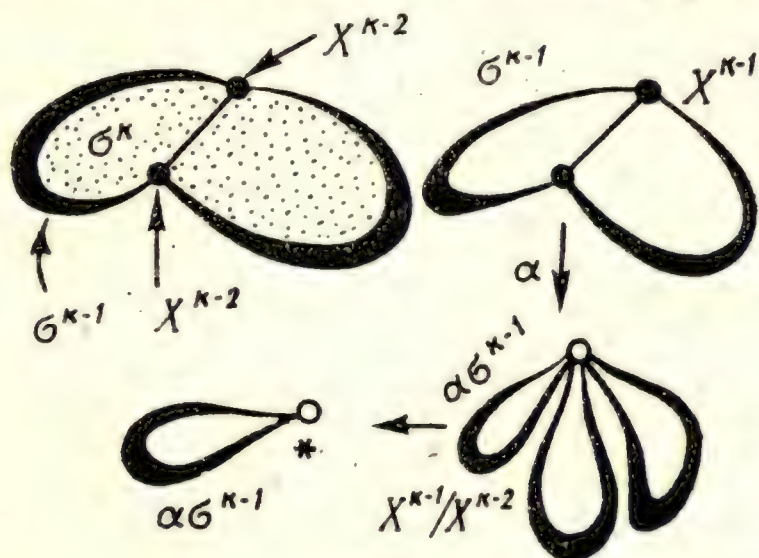


Рис. 14

Мы сопоставили каждой паре клеток σ^k и σ^{k-1} некоторое целое число, называемое коэффициентом инцидентности клеток и обозначаемое обычно так: $[\sigma^k : \sigma^{k-1}]$. Из его определения видно, что это число зависит от выбранных нами ориентаций клеток, при изменении одной из ориентаций число $[\sigma^k : \sigma^{k-1}]$ меняет знак.

Теорема 4. Пусть σ^k — произвольная образующая группы $P_k(X) = H_k(X^k, X^{k-1})$. Тогда действие граничного оператора ∂ на этой образующей задается формулой $\partial \sigma^k = \sum [\sigma^k : \sigma^{k-1}] \sigma^{k-1}$, где сумма берется по всем $(k-1)$ -мерным клеткам σ^{k-1} комплекса X .

Это утверждение дает нам наглядное геометрическое истолкование граничного оператора ∂ , введенного выше на алгебраическом языке для групп клеточных цепей. Если клетка σ^{k-1} не пересекается с замыканием клетки σ^k , то $[\sigma^k : \sigma^{k-1}] = 0$.

Доказательство теоремы 4. Рассмотрим две тройки: $(D^k, S^{k-1}, \emptyset)$ и (X^k, X^{k-1}, X^{k-2}) и непрерывное отображение $(D^k, S^{k-1}, \emptyset) \rightarrow (X^k, X^{k-1}, X^{k-2})$, где $D^k \rightarrow X^k$ — характеристическое отображение клетки σ^k , а $S^{k-1} \rightarrow X^{k-1}$ — ограничение этого отображения на границу шара. Точные последовательности этих троек можно естественно организовать в коммутативную диаграмму:

$$\begin{array}{ccccccc}
 0 & & Z & & Z & & \\
 \parallel & & \parallel & & \parallel & & \\
 H_k(D^k) & \xrightarrow{i_*} & H_k(D^k, S^{k-1}) & \xrightarrow{j} & H_{k-1}(S^{k-1}) & \rightarrow & 0 \\
 & & \downarrow & & \downarrow \Phi_* & & \\
 & & H_k(X^k, X^{k-1}) & \xrightarrow{\partial} & H_{k-1}(X^{k-1}, X^{k-2}) & & \\
 & & \parallel & & \parallel & & \\
 & & P_k(X) & & P_{k-1}(X) & &
 \end{array}$$

Рассмотрим элемент $1 \in Z = H_k(D^k, S^{k-1})$. При гомоморфизме i_* эта образующая отображается в клеточную цепь $1 \cdot \sigma^k \in P_k(X)$, а после применения ∂ переходит в $1 \cdot \partial \sigma^k$. Проследим перемещение этой же образующей по верхней стороне квадрата. При отображении j элемент 1 переходит в $1 \in Z = H_{k-1}(S^{k-1})$ (поскольку j — изоморфизм). При последующем отображении φ_* образующая группы Z переходит в некоторый элемент группы $Z \oplus \dots \oplus Z = H_{k-1}(X^{k-1}/X^{k-2}) = H_{k-1}(\bigvee S^{k-1})$. Каждой образующей группы $P_{k-1}(X) = H_{k-1}(\bigvee S^{k-1})$ соответствует некоторая клетка σ^{k-1} . Ясно, что коэффициент, с которым эта клетка входит в образ единицы при отображении φ_* , в точности равен степени сквозного отображения $S^{k-1} \rightarrow S^{k-1}$, т. е. коэффициенту $[\sigma^k : \sigma^{k-1}]$. Теорема доказана.

Итак, мы получили довольно простое правило вычисления групп сингулярных гомологий клеточного комплекса. Для этого достаточно рассмотреть цепной комплекс клеточных цепей, однозначно определяемый клеточной структурой X , затем в явном виде выписать граничные операторы, для чего достаточно подсчитать коэффициенты инцидентности пар клеток соседних размерностей. После этого нужно вычислить группы гомологий получившегося комплекса групп. Конструкция эта настолько наглядна (а часто и легко вычислима), что она иногда кладется в основу определения групп клеточных гомологий.

Определение 3. Пусть X — конечный клеточный комплекс, $P_k(X)$ — группы клеточных цепей, $\partial_k : P_k(X) \rightarrow P_{k-1}(X)$ — гомоморфизмы, определяемые формулой $\partial_k \sigma^k = \sum [\sigma^k : \sigma^{k-1}] \sigma^{k-1}$. Тогда группы гомологий этого комплекса, т. е. группы $\text{Ker } \partial_k / \text{Im } \partial_{k+1}$ называются группами клеточных гомологий комплекса X .

Для вычисления этих групп не нужно знать сингулярные гомологии X , поскольку все объекты, входящие в определение 2, допускают чисто геометрическое описание (клетки, характеристические отображения, коэффициенты инцидентности). Теорема 3 переформулируется так: группы сингулярных и клеточных гомологий конечного клеточного комплекса изоморфны. Следствие: если одно и то же пространство X представлено двумя способами в виде клеточного комплекса, то группы клеточных гомологий X не зависят от клеточного разбиения, поскольку они изоморфны группам сингулярных гомологий. Таким образом, для вычисления гомологий пространства X мы должны подобрать возможно более простое его представление в виде клеточного комплекса, а затем подсчитать клеточные гомологии.

6. Простейшие примеры вычисления групп клеточных гомологий.

Пример 1. Сфера S^n допускает простейшее клеточное разбиение: $\sigma^0 \cup \sigma^n$, где σ^0 — точка, а σ^n — дополнение к ней в сфере. Ясно, что при $n \geq 1$ имеем: $H_i(S^n) = \mathbb{Z}$ при $i=0, n$ и $H_j(S^n) = 0$ при $j \neq 0, n$.

Пример 2. Вещественное проективное пространство $\mathbf{RP}^n$. Напомним, что одна из его реализаций — это множество последовательностей вида $x = (x_0, x_1, \dots, x_n)$, где x_i — вещественные числа и хотя бы одна координата отлична от нуля, причем последовательность рассматривается с точностью до ненулевого множителя. Простейшее клеточное разбиение $\mathbf{RP}^n$ устроено так: в качестве клетки σ^k нужно взять все последовательности x , для которых $x_k \neq 0$, $x_{k+1} = \dots = x_n = 0$. Тогда в каждой размерности k мы получаем ровно одну клетку σ^k , т. е. $\mathbf{RP}^n = \sigma^0 \cup \sigma^1 \cup \dots \cup \sigma^n$. Следовательно, $P_k(\mathbf{RP}^n) = \mathbf{Z}$. Осталось вычислить граничный оператор $\partial_k: \mathbf{Z} \rightarrow \mathbf{Z}$. На рис. 15 замыкание

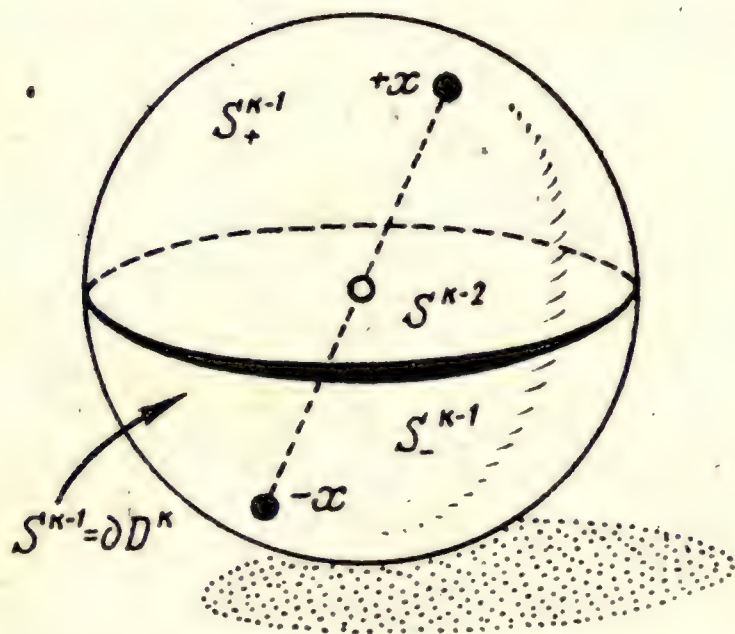


Рис. 15(а)

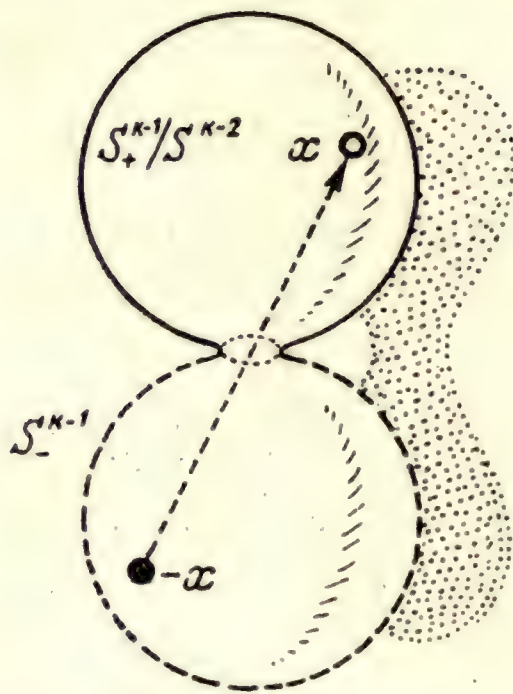


Рис. 15(б)

клетки σ^k , т. е. $\mathbf{RP}^k$ изображено в виде k -мерного шара, граница которого — сфера S^{k-1} — профакторизована по действию группы $\mathbf{Z}_2$, т. е. образующая этой группы представлена преобразованием $x \rightarrow -x$, отражением сферы в начале координат. Другими словами, $\mathbf{RP}^k$ получается из шара D^k при отождествлении на его границе диаметрально противоположных точек. При этом граница шара D^k отображается в фактор $\mathbf{RP}^{k-1}/\mathbf{RP}^{k-2} = S^{k-1}$ следующим образом. Представим S^{k-1} в виде объединения трех непересекающихся подмножеств S_+^{k-1} , S^{k-2} , S_-^{k-1} , где S_+^{k-1} и S_-^{k-1} — открытые полусферы (верхняя и нижняя), а S^{k-2} — экватор. Отображение $\partial D^k = S^{k-1} \rightarrow S^{k-1} = S_+^{k-1}/S^{k-2}$ устроено так: $x \rightarrow x$, если $x \in S_+^{k-1}$; $x \rightarrow *$ при $x \in S^{k-2}$, $-x \rightarrow x$ при $-x \in S_-^{k-1}$. Итак, мы получаем отображение $h: S^{k-1} \rightarrow S^{k-1} = S_+^{k-1}/S^{k-2}$, являющееся диффеоморфизмом на каждом из подмножеств S_+ и S_- . Осталось найти степень этого отображения. Ясно, что прообразом каждой точки $x \in S_+$ являются две точки: она сама и ей диаметрально противоположная на сфере S^{k-1} . Поэтому искомая степень равна либо двум, либо нулю в зависимости от того, меняется или нет ориентация сферы S^{k-1} при отображении $x \rightarrow -x$. Следовательно, нужно найти степень вспомогательного отображения $\alpha: S^{k-1} \rightarrow S^{k-1}$, где $\alpha(x) = -x$.

Лемма 6. Степень отображения $\alpha(x) = -x$ сферы S^{k-1} на себя равна $(-1)^k$.

Отсюда сразу следует, что степень отображения h равна 2 при четном k и равна нулю при нечетном k . Итак, при нечетном k коэффициент $[\sigma^k : \sigma^{k-1}]$ равен нулю, а при четном k — равен 2. Это означает, что граничные операторы в цепном комплексе $\{P_k(X), \partial_k\}$ имеют вид: $\partial \sigma^{2r} = 2\sigma^{2r-1}$, $\partial \sigma^{2r-1} = 0$. Таким образом, мы доказали следующее утверждение.

Предложение 1. Группы сингулярных (и клеточных) гомологий $\mathbb{R}P^n$ имеют вид:

$$\begin{array}{cccccccc} H_0 & H_1 & H_2 & H_3 & H_4 & \dots & H_{n-1} & H_n \\ \mathbb{Z} & \mathbb{Z}_2 & 0 & \mathbb{Z}_2 & 0 & \dots & \mathbb{Z}_2 & 0 & \text{при четном } n; \\ \mathbb{Z} & \mathbb{Z}_2 & 0 & \mathbb{Z}_2 & 0 & \dots & 0 & \mathbb{Z} & \text{при нечетном } n. \end{array}$$

В заключение докажем лемму 6. Рассмотрим сферу S^{k-1} и фиксируем в точке x произвольный касательный орторепер $e(x) = (e_1, \dots, e_{k-1})$. При отображении $\alpha: x \rightarrow -x$ этот репер перейдет в репер $e(-x) = (-e_1, \dots, -e_{k-1})$ (мы считаем, что сфера стандартно вложена в евклидово пространство). Нам нужно сравнить ориентации, индуцированные на сфере этими двумя реперами. Соединим точки x и $-x$ меридианом γ таким, что в точке x его вектором скорости является вектор e_{k-1} (рис. 16). Осуществим гладкую деформацию (перенос) репера $e(x)$ из точки x в точку $-x$, двигаясь по пути γ так, чтобы вектор e_{k-1} оставался все время касательным к γ . Тогда мы получим в точке $-x$ два репера: $-e_1, \dots, -e_{k-2}, -e_{k-1}$ и $-e_1, \dots, -e_{k-2}, +e_{k-1}$. Ясно, что их взаимная ориентация определяется знаком $(-1)^{k-2}$. Лемма доказана.

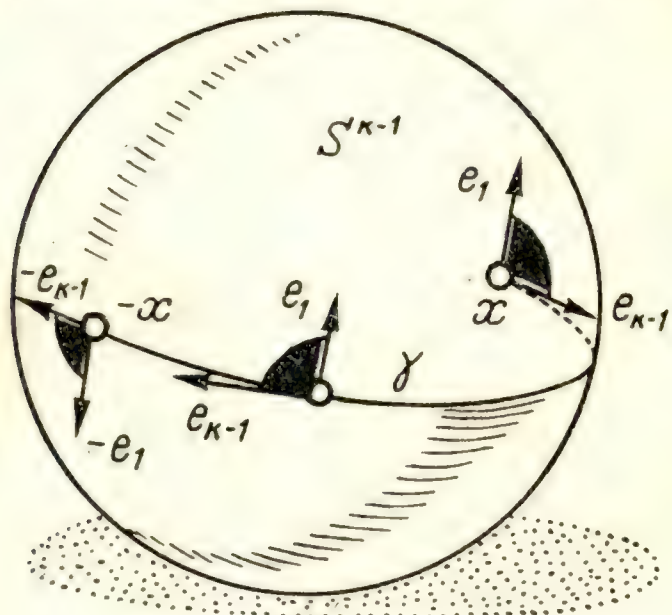


Рис. 16

Тогда мы получим в точке $-x$ два репера: $-e_1, \dots, -e_{k-2}, -e_{k-1}$ и $-e_1, \dots, -e_{k-2}, +e_{k-1}$. Ясно, что их взаимная ориентация определяется знаком $(-1)^{k-2}$. Лемма доказана.

Пример 3. Двумерное компактное связное замкнутое ориентируемое многообразие M_g^2 рода g гомеоморфно двумерной сфере с g ручками и допускает следующее клеточное разбиение: $\sigma^0 \cup \bigcup_{i=1}^{2g} \sigma_i^1 \cup \sigma^2$. Отсюда следует, что гомологии M_g^2 имеют вид: $H_0 = \mathbb{Z}$, $H_1 = \mathbb{Z} \oplus \dots \oplus \mathbb{Z}$ ($2g$ слагаемых), $H_2 = \mathbb{Z}$. Найдите группы гомологий неориентируемого двумерного многообразия.

Определенные выше группы гомологий не исчерпывают собой набор «гомологических инвариантов», позволяющих иногда различать клеточные комплексы. Теорию гомологий можно построить, используя цепи с коэффициентами в произвольной абелевой группе A . Это означает, что нужно рассмат-

ривать цепи как линейные комбинации с коэффициентами из A . Ясно, что все конструкции переносятся на этот случай автоматически, что позволяет определить группы $H_k(X, A)$, называемые группами гомологий с коэффициентами в группе A . Как и для группы $A = \mathbb{Z}$ (рассматривавшейся выше), верны результаты, относящиеся к клеточным гомологиям. В частности, эти группы могут быть определены как гомологии цепного комплекса, составленного из групп клеточных цепей $\{\sum a_i \sigma_i^k, a_i \in A\} = P_k(X, A)$ с коэффициентами в A .

Группы гомологий $H_k(X)$, изучавшиеся нами выше, в новых обозначениях записываются как $H_k(X, \mathbb{Z})$. Для сокращения записи мы будем часто опускать обозначение группы коэффициентов, кроме тех случаев, когда от выбора группы A зависит конечный результат.

Задача. Для примеров 1—3 (см. выше) вычислить группы гомологий с коэффициентами в группах: $A = \mathbb{R}, \mathbb{Q}, \mathbb{Z}_2, \mathbb{Z}_p$, где $\mathbb{R}$ — поле вещественных чисел, $\mathbb{Q}$ — поле рациональных чисел, p — простое.

§ 2. Группы когомологий и препятствия к продолжению отображений

1. Сингулярные коцепи и оператор δ .

Пусть X — клеточный комплекс и $C_k(X)$ — группа k -мерных сингулярных цепей пространства X . Пусть A — абелева группа.

Определение 1. Сингулярной коцепью пространства X с коэффициентами в группе A называется гомоморфизм группы $C_k(X)$ в группу A . Естественная операция сложения коцепей превращает множество коцепей в абелеву группу, обозначаемую $C^k(X, A)$ и называемую группой коцепей пространства X .

Рассмотрим граничный оператор $\partial : C_k(X) \rightarrow C_{k-1}(X)$. Пусть $h \in C^{k-1}(X, A)$ — произвольная коцепь, т. е. гомоморфизм $h : C_{k-1}(X) \rightarrow A$. Тогда однозначно определяется коцепь $\delta h \in C^k(X, A)$, задаваемая формулой: $\delta h(\alpha) = h(\partial \alpha)$, т. е. $\delta h : C_k(X) \rightarrow A$. Другими словами, коцепь δh определяется из диаграммы:

$$\begin{array}{ccc} \alpha \in C_k(X) & \xrightarrow{\partial} & C_{k-1}(X) \\ \downarrow \delta h & & \downarrow h \\ & \xrightarrow{\quad} A & \xleftarrow{\quad} \end{array}$$

Иногда будем обозначать оператор $\delta : C^{k-1}(X, A) \rightarrow C^k(X, A)$ через δ_{k-1} .

Определение 2. Оператор $\delta_{k-1} : C^{k-1} \rightarrow C^k$ называется кограничным. Он сопряжен к граничному оператору ∂ .

Поскольку $\partial^2 = 0$, то $\delta^2 = 0$. Следовательно, мы получаем последовательность групп и связывающих их гомоморфизмов сле-

дующего вида:

$$C^0(X, A) \xrightarrow{\delta_0} C^1(X, A) \xrightarrow{\delta_1} C^2(X, A) \rightarrow \dots,$$

где $\delta_k \delta_{k-1} = 0$. Эта последовательность называется коцепным комплексом. Поступая по схеме предыдущего параграфа, рассмотрим группы $\text{Ker } \delta$ и $\text{Im } \delta$, после чего построим группу $H^k(X, A) = \text{Ker } \delta_k / \text{Im } \delta_{k-1}$.

Определение 3. Группы $H^k(X, A)$ называются группами когомологий пространства X с коэффициентами в абелевой группе A . Элементы группы $B^k = \text{Im } \delta_{k-1}$ называются кограницами, а элементы группы $Z^k = \text{Ker } \delta_k$ называются коциклами.

Если пространство X линейно связно, то $H^0(X, A) = A$. Как и в случае гомологий, естественно определяются группы относительных когомологий. Пусть Y — замкнутый подкомплекс в комплексе X , тогда $C_k(Y) \subset C_k(X)$. Пусть $C^k(X, Y)$ — группа всех тех гомоморфизмов $\alpha: C_k(X) \rightarrow A$, которые равны нулю на $C_k(Y)$. Ясно, что $\delta C^k(X, Y) \subset C^{k+1}(X, Y)$, а потому возникают группы $H^k(X, Y, A) = \text{Ker } \delta / \text{Im } \delta$, которые называются группами относительных когомологий. Как и в случае гомологий, естественно возникает точная последовательность пары и тройки. Опуская детали построения, приведем только точную последовательность пары (проверьте точность!):

$$\dots \rightarrow H^k(X, Y, A) \rightarrow H^k(X, A) \rightarrow H^k(Y, A) \xrightarrow{\delta} H^{k+1}(X, Y, A) \rightarrow \dots$$

и тройки (X, Y, Z) :

$$\dots \rightarrow H^k(X, Y, A) \rightarrow H^k(X, Z, A) \rightarrow H^k(Y, Z, A) \xrightarrow{\delta} H^{k+1}(X, Y, A) \rightarrow \dots$$

Группы сингулярных когомологий гомотопически инвариантны. Поступая по схеме § 1, определим группы клеточных когомологий. Для этого вводятся группы клеточных коцепей $P^k(X, A)$, определяемые как группы $H^k(X^k, X^{k-1}, A)$, где X^k — k -мерный остов клеточного комплекса X . Из точной последовательности тройки (X^{k+1}, X^k, X^{k-1}) возникает кограничный оператор $\delta: P^k(X, A) \rightarrow P^{k+1}(X, A)$. Группы клеточных когомологий определяются как группы $\text{Ker } \delta / \text{Im } \delta$ для коцепного комплекса $\{P^k(X, A), \delta\}$.

Теорема 1. Для конечного клеточного комплекса X группы сингулярных когомологий и группы клеточных когомологий изоморфны и являются конечно-порожденными абелевыми группами, если таковой является группа коэффициентов.

Доказательство проводится так же, как в § 1.

2. Задача продолжения непрерывного отображения с подпространства на все пространство.

При изучении топологических вариационных задач (чему посвящены последующие главы) нам придется иногда решать следующую задачу.

Пусть Y — замкнутое подпространство топологического пространства X и пусть задано непрерывное отображение этого подпространства в некоторое пространство Z . Возникает вопрос: в каких случаях это отображение можно продолжить до непрерывного отображения всего пространства X в пространство Z ? Ясно, что такое непрерывное продолжение существует не всегда. Имеются топологические препятствия, не позволяющие в некоторых случаях распространять отображение с подпространства на все пространство. Изучением всех разнообразных версий этого вопроса занимается теория препятствий. Здесь мы приведем лишь те результаты, которые будут использоваться нами в дальнейшем. Мы уже располагаем всем необходимым материалом, разработанным в предыдущих параграфах, чтобы изложить необходимые для нас сведения.

3. Препятствие к продолжению отображения.

Рассмотрим конечный клеточный комплекс K и топологическое пространство X . Пусть задано непрерывное отображение g $(n-1)$ -мерного остова K^{n-1} клеточного комплекса K (т. е. объединения всех его клеток размерностей, не превосходящих $n-1$) в пространство X . Мы хотим продолжить его до отображения f следующего остова K^n в X . Для упрощения рассуждений будем считать, что X либо односвязно, либо, в том случае, когда X одномерно, имеет коммутативную фундаментальную группу. Рассмотрим все n -мерные клетки σ^n , входящие в n -мерный остов комплекса K . В силу его конечности этих клеток конечное число. Для построения искомого продолжения надо уметь продолжать исходное отображение на каждую n -мерную клетку по отдельности. Фиксируем какую-либо клетку σ^n . Ее граница уже отображена при помощи отображения g в пространство X . Нужно продолжить это отображение на всю клетку. Будем считать, что конструируемое отображение $f: K^n \rightarrow X$ совпадает с отображением g на остова $K^{n-1} \subset K^n$. Пусть $\chi: D^n \rightarrow K$ — характеристическое отображение клетки σ^n , где D^n — n -мерный диск. Возникает сквозное отображение $S^{n-1} = \partial D^n \xrightarrow{\chi} K^{n-1} \xrightarrow{g} X$. Здесь мы используем то, что граница клетки σ^n содержится в остова K^{n-1} согласно определению клеточного комплекса. Полученное отображение $g\chi: S^{n-1} \rightarrow X$ определяет некоторый элемент $[g\chi]$ гомотопической группы $\pi_{n-1}(X)$. Это определение корректно в силу наложенных на пространство X ограничений. Они гарантируют нам, что определение элементов группы $\pi_{n-1}(X)$ не зависит от выбора базисной (отмеченной) точки. В общем случае нужно требовать, чтобы пространство X было $(n-1)$ -простым (см. определение в [4]), т. е. чтобы фундаментальная группа тривиально (с гомотопической точки зрения) действовала на $\pi_{n-1}(X)$. Это действие см. на рис. 17. Ясно, что в указанных выше предположениях X является $(n-1)$ -простым.

Отображение $g\chi: S^{n-1} \rightarrow X$ можно продолжить до отображе-

ния всего диска D^n в том и только в том случае, когда элемент $[g\chi]$ из $\pi_{n-1}(X)$ равен нулю. В общем случае мы сопоставили каждой клетке σ^n некоторый элемент $[g\chi]$ абелевой группы $G = \pi_{n-1}(X)$. Рассмотрим группу $P_n(K)$ клеточных цепей комплекса K . Ясно, что построенное соответствие $\sigma^n \rightarrow [g\chi] \in G$ можно естественно

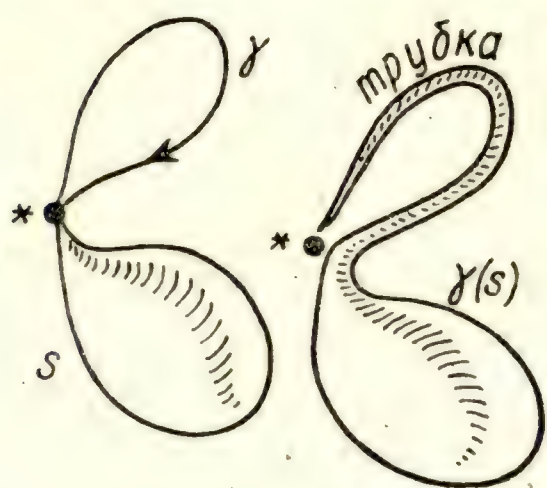


Рис. 17

доопределить до некоторого гомоморфизма группы цепей $P_n(K)$ в группу G . Для этого достаточно построить отображение для каждой клетки и продолжить на все цепи по линейности. Тем самым мы определяем некоторую коцепь c_g^n из группы клеточных коцепей.

Определение 4. Коцепь $c_g^n \in P^n(K)$ с коэффициентами из абелевой группы $G = \pi_{n-1}(X)$ называется препятствием к продолжению отображения g с остова K^{n-1} на остов K^n .

Лемма 1. Отображение $g: K^{n-1} \rightarrow X$ может быть продолжено до непрерывного отображения $f: K^n \rightarrow X$ тогда и только тогда, когда коцепь c_g^n тождественно равна нулю.

Доказательство следует из определения препятствия.

Лемма 2. Пусть пространство X $(n-1)$ -просто. Коцепь $c_g^n \in P^n(K, \pi_{n-1}(X))$ является коциклом, т. е. $\delta c_g^n = 0$. Следовательно, коцепь c_g^n определяет некоторый класс когомологий $C_g^n = [c_g^n]$ из группы $H^n(K, \pi_{n-1}(X))$.

Так как идеи, лежащие в основе доказательства, нами не будут в дальнейшем использоваться, то доказательство мы опускаем (см. [4]). В тех приложениях, с которыми мы столкнемся, равенство $\delta c_g^n = 0$ будет следовать из наглядных соображений.

Теорема 2. Пусть пространство X $(n-1)$ -просто. Коцикл C_g^n равен нулю (как элемент группы когомологий) тогда и только тогда, когда исходное отображение $g: K^{n-1} \rightarrow X$ можно продолжить до непрерывного отображения $f: K^n \rightarrow X$, предварительно изменив отображение g на остова K^{n-1} и не меняя его на остова K^{n-2} .

В отличие от леммы 2 схема доказательства теоремы 2 будет использоваться в приложениях, поэтому мы приведем доказательство.

Доказательство. Для построения отображения f нам потребуется новое понятие, тесно связанное с препятствием c_g^n , а именно понятие различающей коцепи. Пусть даны два отображения: $f, g: K^{n-1} \rightarrow X$, совпадающие на остова K^{n-2} , т. е. $f(x) = g(x)$ для $x \in K^{n-2}$. Пусть $\sigma^{n-1} \subset K^{n-1}$ — произвольная клетка и $\chi: D^{n-1} \rightarrow K$ — ее характеристическое отображение. Так как граница S^{n-2} диска D^{n-1} отображается в остов K^{n-2} , то два сквозных отображения $f\chi$ и $g\chi$ одинаково отображают сферу S^{n-2} в комплекс K . Клетка σ^{n-1} отображается, вообще говоря, по-разному, однако эти два образа имеют общую гра-

ницу, будучи склеены по образу сферы S^{n-2} в K . Следовательно, мы получаем в пространстве X сфероид, определяющий некоторый элемент группы $\pi_{n-1}(X)$ и измеряющий отклонение отображения f от отображения g на σ^{n-1} . Тем самым мы сопоставили каждой клетке σ^{n-1} элемент группы $\pi_{n-1}(X)$ (здесь использована $(n-1)$ -простота X). Получаем некоторый гомоморфизм группы цепей $P_{n-1}(K)$ в группу $\pi_{n-1}(X)$, т. е. получаем $(n-1)$ -мерную коцепь. Эта коцепь обозначается через $d_{f,g}^{n-1}$ и называется различающей коцепью двух отображений f и g (рис. 18). Из ее определения следует, что $d_{f,g}^{n-1} = 0$ тогда и только тогда, когда существует гомотопия, соединяющая f и g и постоянная на остове K^{n-2} , где f и g совпадают.

Лемма 3. Для любого отображения $f: K^{n-1} \rightarrow X$ и любой коцепи $d \in P^{n-1}(K, \pi_{n-1}(X))$ всегда можно подобрать такое отображение $g: K^{n-1} \rightarrow X$, которое совпадает с f на остове K^{n-2} и коцепь d является различающей для f и g , т. е. $d = d_{f,g}^{n-1}$.

Доказательство. Рассмотрим произвольную клетку σ^{n-1} из K и ее образ при отображении f в X . Отметим в цент-

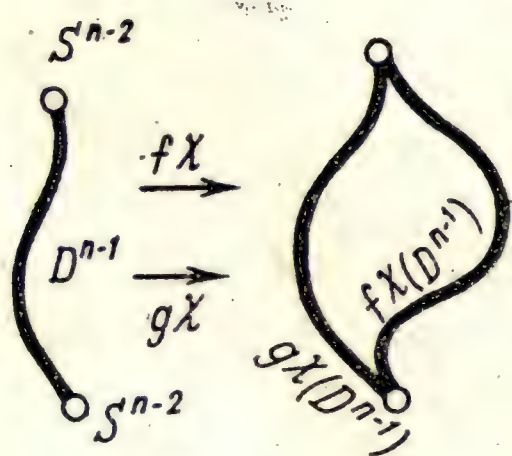


Рис. 18

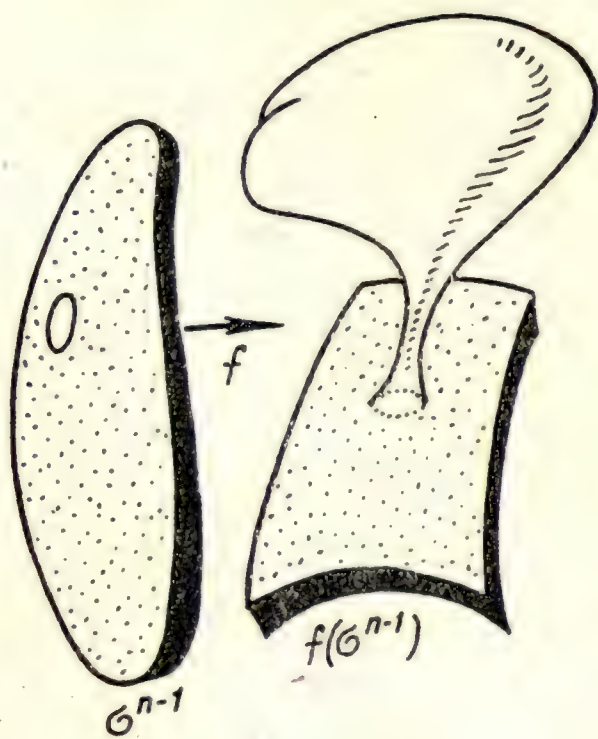


Рис. 19

ре клетки достаточно малый шар и, рассмотрев его образ при f в X , вырежем этот образ (используя теорему о клеточной аппроксимации [4]) из образа клетки σ^{n-1} . Затем к полученной дырке приклеим сфероид, являющийся представителем того элемента из группы $\pi_{n-1}(X)$, который является значением коцепи d на клетке σ^{n-1} (рис. 19). В качестве нового отображения g возьмем отображение, совпадающее с f всюду, кроме отмеченного шара, а на шаре совпадающее с отображением, реализующим указанный сфероид. Грубо говоря, отмеченный маленький шар выпучивается в сфероид, реализующий значение коцепи d на этой клетке. Выполняя эту операцию на каждой клетке, мы и получаем некоторое отображение g . Сравнивая

его с исходным, очевидно, получаем, что $d = d_{f,g}^{n-1}$. Лемма доказана.

Лемма 4. Имеет место равенство: $\delta d_{f,g}^{n-1} = c_f^n - c_g^n$.

Доказательство. В силу определения кограничного оператора в терминах клеточных коцепей и цепей имеем равенство:

$$(\delta d_{f,g}^{n-1}) \sigma^n = d_{f,g}^{n-1} (\partial \sigma^n) = \sum [\sigma^n : \sigma_i^{n-1}] d_{f,g}^{n-1} (\sigma_i^{n-1}).$$

Для подсчета коэффициентов инцидентности нужно рассмотреть следующее сквозное отображение

$$S^{n-1} = \partial D^n \xrightarrow{\chi} K^{n-1} \xrightarrow{\pi} K^{n-1}/K^{n-2} = \bigvee_{p_i} S_i^{n-1} \xrightarrow{p_i} S_i^{n-1},$$

где π — естественная факторизация, а p_i — проекция букета сфер на сферу с номером i в этом букете. Степень полученного отображения $S^{n-1} \rightarrow S_i^{n-1}$ и называется коэффициентом инцидентности $[\sigma^n : \sigma_i^{n-1}]$. Легко доказать, что отображение $S^{n-1} \rightarrow K^{n-1}$ гомотопно такому отображению, при котором почти вся

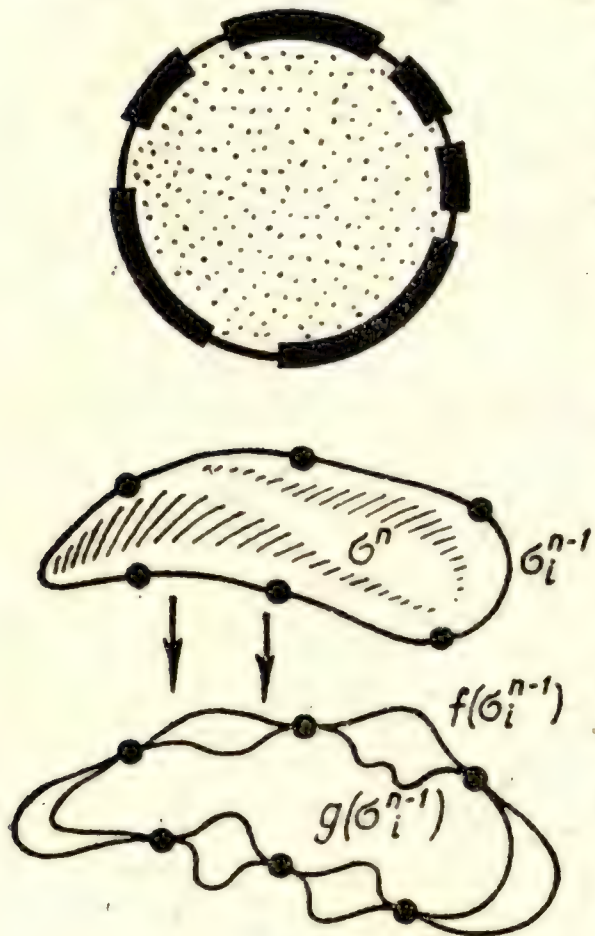


Рис. 20

сфера S^{n-1} , за исключением конечного числа шариков на этой сфере, отображается в остов K^{n-2} , а шарики отображаются в остов K^{n-1} , причем каждый шарик отображается на свою клетку σ_i^{n-1} со степенью ± 1 . Алгебраическое число шариков, отобразившихся на клетку σ_i^{n-1} (т. е. сумма степеней ± 1), в точности равно искомому коэффициенту инцидентности (рис. 20). На этом рисунке толстые отрезки на границе шара условно изображают маленькие шарики размерности $n-1$, отображающиеся на клетки σ_i^{n-1} . Подсчитаем, какое значение принимает коцепь $c_f^n - c_g^n$ на клетке σ^n (рис. 20), т. е. куда отобразятся граничные клетки σ_i^{n-1} . В итоге нам нужно выяснить, куда отобразятся шарики, отмеченные на S^{n-1} , при сквозном отображении с помощью f и g . Так

как отображения f и g совпадают на остоле K^{n-2} , то для подсчета значения коцепи на клетке σ^n нужно просуммировать по границе $\partial \sigma^n$, т. е. по всем клеткам σ_i^{n-1} , следующие сфероиды из $\pi_{n-1}(X)$. На каждой клетке σ_i^{n-1} нужно взять сфероид, совпадающий со значением различающей коцепи $d_{f,g}^{n-1}$ на клетке σ_i^{n-1} , причем этот сфероид нужно взять столько раз, сколько на клетку σ_i^{n-1} отобразится маленьких шариков. Очевидно, что эта сумма дает нам коэффициент инцидентности, а всей клетке σ^n будет, следовательно,

сопоставлена сумма $\sum [\sigma^n : \sigma_i^{n-1}] d_{f,g}^{n-1}(\sigma_i^{n-1})$, совпадающая с $(\delta d_{f,g}^{n-1}) \sigma^n$, что и завершает доказательство.

Вернемся к доказательству теоремы 2. Пусть отображение $g : K^{n-1} \rightarrow X$ продолжается до отображения $f : K^n \rightarrow X$ без изменения на остове K^{n-2} , но, возможно, с изменением на остове K^{n-1} . В силу леммы 4 имеем: $\delta d_{f,g}^{n-1} = c_f^n - c_g^n$. Но так как f определено на K^n , то $c_f^n = 0$, следовательно, $c_g^n = -\delta d_{f,g}^{n-1}$, и мы получаем, что $C_g^n = 0$. Докажем обратное утверждение. Пусть дано, что $C_g^n = 0$, т. е. $c_g^n = \delta d$, где d — некоторая коцепь. Согласно лемме 3 существует отображение $f : K^{n-1} \rightarrow X$, совпадающее с отображением g на остове K^{n-2} и такое, что $-d = d_{f,g}^{n-1}$. Тогда в силу леммы 4 получаем, что $c_f^n = c_g^n + \delta d_{f,g}^{n-1} = c_g^n - \delta d = 0$. Следовательно, отображение f продолжается до отображения $K^n \rightarrow X$. Таким образом, мы продолжили отображение g , предварительно изменив его, быть может, на остове K^{n-1} , но не на K^{n-2} , что и требовалось доказать.

4. Случай существования ретракции пространства на подпространство, гомеоморфное сфере.

Докажем теорему о ретракции, которая будет нужна при изучении вариационных задач. Рассмотрим в качестве группы коэффициентов теории клеточных гомологий абелеву группу (по сложению) $U = \mathbf{R}^1(\text{mod } 1)$, т. е. окружность.

Теорема 3 (Хопф). Пусть задано вложение i сферы S^{n-1} в конечный n -мерный клеточный комплекс K . Пусть гомоморфизм $i_* : H_{n-1}(S^{n-1}, U) \rightarrow H_{n-1}(K, U)$, индуцированный этим вложением, является мономорфизмом. Тогда сфера $S_0^{n-1} = iS^{n-1}$ является ретрактом K , т. е. существует непрерывное отображение $f : K \rightarrow S_0^{n-1}$, тождественное на сфере S_0^{n-1} .

Доказательство. Приводимый нами конструктивный вариант доказательства, необходимый для явного построения конкретных ретракций, принадлежит Т. Н. Фоменко. Нам нужно построить встречное непрерывное отображение $f : K \rightarrow S^{n-1}$, причём такое, чтобы его ограничение на сферу iS^{n-1} , вложенную в K , было тождественным отображением сферы iS^{n-1} на S^{n-1} . Для построения такого отображения применим теорию препятствий. Прежде всего выберем для сферы S^{n-1} простейшее клеточное разбиение в сумму двух клеток: нульмерной $*$ и $(n-1)$ -мерной σ^{n-1} . Следовательно, мы можем конструировать искомое отображение так, чтобы оно было клеточным и переводило весь остов K^{n-2} в точку на сфере S^{n-1} . Другими словами, достаточно построить непрерывное отображение факторкомплекса K/K^{n-2} в сферу S^{n-1} ; здесь $K = K^n$. Для применимости теории препятствий нужно в качестве X рассматривать $(n-1)$ -простое пространство. Возьмем $X = S^{n-1}$, тогда условие $(n-1)$ -простоты, очевидно, выполнено. В самом деле, если $n-1 > 1$, то $\pi_1(X) = 0$, а если $n-1 = 1$, то группа $\pi_1(S^1) = \mathbf{Z}$ абелева и действует на себе тождественно. Так как остов K^{n-2} стянут нами в точку, то можно считать, что остов K^{n-1} совпа-

дает с букетом $(n-1)$ -мерных сфер $S_0^{n-1} \vee S_1^{n-1} \vee \dots \vee S_k^{n-1}$. Ясно, что сферу iS^{n-1} можно считать одной из них. Пусть для определенности это будет сфера S_0^{n-1} . Каждую из сфер S_i^{n-1} представим в виде простейшего клеточного разбиения и обозначим соответствующие $(n-1)$ -мерные клетки через σ_i^{n-1} . Мы должны построить непрерывное отображение $f: K \rightarrow S^{n-1}$ такое, чтобы его ограничение на S_0^{n-1} было тождественным отображением на S^{n-1} . Пусть $\sigma_1^n, \dots, \sigma_q^n$ — n -мерные клетки комплекса K . Тогда можно считать, что $K = \sigma^0 \cup \sigma_0^{n-1} \cup \dots \cup \sigma_k^{n-1} \cup \sigma_1^n \cup \dots \cup \sigma_q^n$ (рис. 21).

Предварительно мы должны построить некоторое отображение остова K^{n-1} в S^{n-1} . Сейчас мы построим некоторое такое

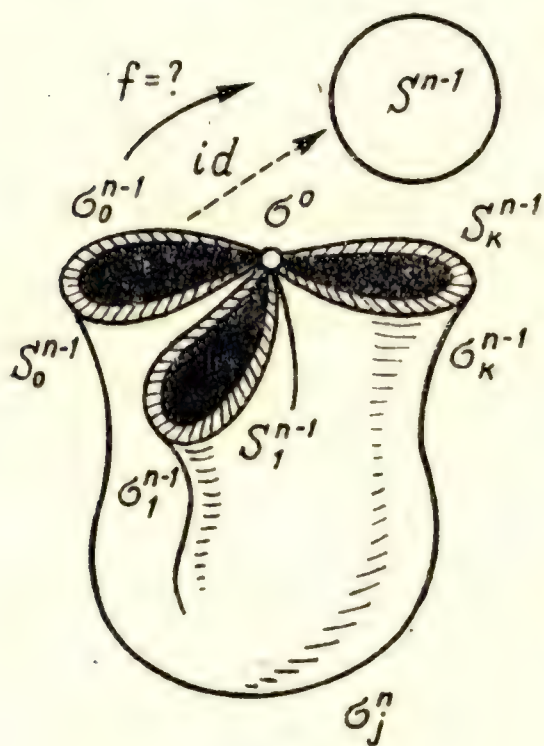


Рис. 21

отображение g_1 , которое, вообще говоря, не будет продолжаться до отображения всего K в S^{n-1} , однако может быть использовано для построения второго, уже окончательного отображения g_2 , продолжающегося до отображения $f: K \rightarrow S^{n-1}$. Так как остов K^{n-1} является букетом сфер, то достаточно задать отображение g_1 на каждой из них. На выделенной сфере S_0^{n-1} мы зададим тождественное отображение $S_0^{n-1} \rightarrow S^{n-1}$, а все остальные сферы отобразим в отмеченную точку $* \in S^{n-1}$. Это отображение непрерывно, однако легко построить примеры, показывающие, что оно не обязано продолжаться до непрерывного отображения всего комплекса в сферу. Подсчитаем препятствие $c_{g_1}^n$ для отображения g_1 .

Рассмотрим произвольную n -мерную клетку σ^n . Для подсчета $c_{g_1}^n(\sigma^n)$ нужно рассмотреть сквозное отображение $(n-1)$ -мерной сферы — границы клетки — в сферу S^{n-1} и найти его степень. Здесь мы используем то, что $\pi_{n-1}(S^{n-1}) = \mathbb{Z}$ и гомотопические классы отображений сферы в сферу классифицируются степенью отображения. Это сквозное отображение является композицией характеристического отображения и отображения $g_1: K^{n-1} \rightarrow S^{n-1}$. Та часть граничной сферы $\partial\sigma^n$, которая отобразилась в букет $S_1^{n-1} \vee \dots \vee S_k^{n-1}$, т. е. неинцидентна с выделенной сферой S_0^{n-1} , отобразится в итоге в отмеченную точку $*$ на сфере S^{n-1} . Следовательно, эти участки сферы $\partial\sigma^n$ не принимают участия в формировании искомой степени (вклад равен нулю). Остальная часть граничной сферы $\partial\sigma^n$ отобразится на выделенную сферу S_0^{n-1} , «намотается» на нее столько раз, каков коэффициент инцидентности $[\sigma^n: \sigma_0^{n-1}]$, а затем с помощью тождественного отображения отобразится на сферу S^{n-1} . В результате граничная сфера $\partial\sigma^n$ отображается на сферу S^{n-1} со степенью, равной коэффициенту инцидентности

$[\sigma^n : \sigma_0^{n-1}]$. Итак, мы подсчитали, что $c_{g_1}^n(\sigma^n) = [\sigma^n : \sigma_0^{n-1}]$. Мы утверждаем, что эта коцепь представима в виде кограницы от некоторой $(n-1)$ -мерной коцепи d .

Лемма 5. В предположениях теоремы 3 коцепь $c_{g_1}^n$ является кограницей, т. е. $c_{g_1}^n = \delta d$, где $d \in P^{n-1}(K, \pi_{n-1}(S^{n-1}))$ и $d(\lambda \sigma_0^{n-1}) = 0$, $0 \leq \lambda \leq 1$.

Доказательство. Рассмотрим для простоты тот случай, когда среди n -мерных клеток комплекса K имеется всего лишь одна клетка σ^n , имеющая ненулевой коэффициент инцидентности с выделенной клеткой σ_0^{n-1} . В грубых чертах дальнейшие рассуждения будут развиваться так. Наша цель — построить $(n-1)$ -мерную коцепь d на букете сфер, причем эта коцепь должна равняться нулю на выделенной клетке σ_0^{n-1} (более точно, на элементарной цепи $1 \cdot \sigma_0^{n-1}$) и быть, вообще говоря, нетривиальной на остальных сферах букета, т. е. мы хотим «снять» коцепь с выделенной сферы S_0^{n-1} . Воспользуемся тем, что цикл S_0^{n-1} не является гомологичным нулю в комплексе K для группы коэффициентов U . Напомним, что гомоморфизм, индуцированный вложением сферы, не имеет ядра. Вычислим границу клетки σ^n . Имеем: $\partial \sigma^n = [\sigma^n : \sigma_0^{n-1}] \sigma_0^{n-1} + \dots + [\sigma^n : \sigma_k^{n-1}] \sigma_k^{n-1}$. Обозначая $[\sigma^n : \sigma_i^{n-1}]$ через a_i , можем записать: $\partial \sigma^n = a_0 \sigma_0^{n-1} + a_1 \sigma_1^{n-1} + \dots + a_k \sigma_k^{n-1}$. Таким образом, цепь $-a_0 \sigma_0^{n-1}$ гомологична цепи $a_1 \sigma_1^{n-1} + \dots + a_k \sigma_k^{n-1}$. Следовательно, определяемые ими циклы совпадают как элементы группы гомологий $H_{n-1}(K, \mathbb{Z})$.

Из условия теоремы вытекает, что цикл $\lambda \sigma_0^{n-1}$, где $0 < \lambda < 1$, отличен от нуля в группе $H_{n-1}(X, U)$. Рассмотрим следующие три случая: а) все целые числа $a_1, \dots, a_k$ взаимно-просты; б) числа $a_1, \dots, a_k$ имеют наибольший общий делитель p , отличный от единицы, причем число a_0/p не является целым, т. е. p не делит a_0 ; в) числа $a_1, \dots, a_k$ имеют наибольший общий делитель p , отличный от единицы, и p делит a_0 .

Рассмотрим случай а).

В силу взаимной простоты чисел $a_1, \dots, a_k$ существуют целые числа $x_1, \dots, x_k$ такие, что $a_0 = x_1 a_1 + \dots + x_k a_k$. Определим $(n-1)$ -мерную коцепь d на остоле K^{n-1} так: $d(\lambda \sigma_i^{n-1}) = \lambda x_i$, $x_i \in \mathbb{Z} = \pi_{n-1}(S^{n-1})$, $1 \leq i \leq k$ и $d(\lambda \sigma_0^{n-1}) = 0$. Здесь числа $x_1, \dots, x_k$ реализуются как степени отображений сфер S_i^{n-1} на сферу S^{n-1} . Ясно, что $\delta d = c_{g_1}^n$. В самом деле:

$$(\delta d) \sigma^n = d(\partial \sigma^n) = \sum_{i=1}^k a_i d \sigma_i^{n-1} = \sum_{i=1}^k x_i a_i = a_0 = c_{g_1}^n(\sigma^n).$$

Итак, в случае а) лемма доказана.

В случае б) рассмотрим равенство: $\partial(\lambda \sigma^n) - \lambda a_0 \sigma_0^{n-1} = \lambda p (b_1 \sigma_1^{n-1} + \dots + b_k \sigma_k^{n-1})$, где $a_i = p b_i$, $1 \leq i \leq k$. Здесь числа $b_1, \dots, b_k$ взаимно-простые. Так как число a_0/p не целое, то, используя группу коэффициентов U , получаем, что существует не-

нулевой цикл $\frac{a_0}{p}\sigma_0^{n-1}$ (при $\lambda=1/p$) в гомологиях $H_{n-1}(S^{n-1}, U)$, который становится гомологичным нулю в группе $H_{n-1}(K, U)$, так как $\frac{a_0}{p}\sigma_0^{n-1} \equiv \partial\left(\frac{1}{p}\sigma^n\right) \pmod{1}$. Это противоречит мономорфности гомоморфизма $H_{n-1}(S^{n-1}, U) \rightarrow H_{n-1}(K, U)$. Следовательно, случай б) в действительности не реализуется.

Наконец, рассмотрим случай в). Так как p — наибольший общий делитель чисел $a_1, \dots, a_k$, то $a_i = pb_i$, $1 \leq i \leq k$, $a_0 = mp$, $m \in \mathbb{Z}$ и числа $b_1, \dots, b_k$ взаимно-простые. Следовательно, существуют целые числа $x_1, \dots, x_k$ такие, что $m = x_1b_1 + \dots + x_kb_k$. Реализуем числа x_i как степени отображений сфер S_i^{n-1} в сферу S^{n-1} и определим коцепь d следующим образом: $d(\sigma_i^{n-1}) = x_i \in \pi_{n-1}(S^{n-1})$, $d(\sigma_0^{n-1}) = 0$. Утверждается, что $\delta d = c_{g_1}^n$. В самом деле, $(\delta d)\sigma^n = d(\partial\sigma^n) = d(a_0\sigma_0^{n-1} + \dots + a_k\sigma_k^{n-1}) = a_1x_1 + \dots + a_kx_k = mp = a_0 = c_{g_1}^n(\sigma^n)$. Лемма доказана.

Возвращаемся к доказательству теоремы 3. В силу леммы 3 для отображения $g_1: K^{n-1} \rightarrow S^{n-1}$ и для коцепи d , построенной выше, т. е. такой, что $\delta d = c_{g_1}^n$, можно подобрать такое непрерывное отображение $g_2: K^{n-1} \rightarrow S^{n-1}$, что выполнено равенство $d = d_{g_1, g_2}^{n-1}$. В силу леммы 4 имеем: $\delta d_{g_1, g_2}^{n-1} = \delta d = c_{g_1}^n - c_{g_2}^n$, т. е. $c_{g_2}^n = 0$. Итак, мы нашли отображение g_2 $(n-1)$ -мерного остова K в S^{n-1} , которое продолжается до непрерывного отображения $f: K^n \rightarrow S^{n-1}$. Осталось убедиться в том, что это отображение f тождественно на сфере S_0^{n-1} . Это следует из доказательства леммы 3. В самом деле, отображение g_2 строится так. Оно отличается от g_1 лишь на тех $(n-1)$ -мерных клетках, на которых коцепь d отлична от нуля. Но коцепь d была специально построена так, что она равна нулю на цепях вида $\lambda\sigma_0^{n-1}$, следовательно, g_2 , как и g_1 , является тождественным отображением сферы S_0^{n-1} на сферу S^{n-1} . Если имеется несколько n -мерных клеток, инцидентных с клеткой σ_0^{n-1} , то доказательство проводится по той же схеме, и мы его предоставляем читателю. Теорема доказана.

Приведенное доказательство, несмотря на геометрическую простоту, довольно громоздко. Мы избрали этот путь, поскольку он окажется для нас полезным в дальнейшем, когда нам потребуется строить конкретную ретракцию некоторого комплекса на сферу. Явное построение искомой ретракции становится элементарным после проведенных выше рассуждений. Однако нам представляется уместным привести здесь и другое, более короткое доказательство теоремы 3, может быть не столь наглядное, зато проясняющее общие топологические причины, влекущие за собой существование искомой ретракции. Правда, нам придется воспользоваться свойствами так называемых комплексов Эйленберга — Маклейна. Фиксируем натуральное число $n \geq 1$ и группу π , которую будем предполагать абелевой при $n > 1$. Тогда существует клеточный комплекс, обозначаемый

через $K(\pi, n)$ и называемый пространством Эйленберга — Мак-лейна, такой, что все его гомотопические группы тривиальны, кроме группы $\pi_n(K(\pi, n))$, изоморфной группе π . Имеет место важное утверждение: множество $\Pi(K, K(\pi, n))$ гомотопических классов отображений конечного клеточного комплекса K в пространство $K(\pi, n)$ находится во взаимно-однозначном соответствии с множеством элементов группы когомологий $H^n(K, \pi)$. Доказательство см., например, в [4]. Теорема 3 выводится отсюда так. Рассмотрим сферу S^{n-1} и вложим ее в пространство $K(\mathbb{Z}, n-1)$ в виде $(n-1)$ -мерного остова. Для этого нужно приклеить к сфере некоторое количество клеток, чтобы «уничтожить» все ее гомотопические группы, начиная с размерности n . При этом мы пользуемся тем, что $\pi_{n-1}(S^{n-1}) = \mathbb{Z}$. Первой гомотопической группой, нуждающейся в аннулировании, является группа $\pi_n(S^{n-1})$. Следовательно, мы должны приклеить к сфере S^{n-1} клетки размерностей $n+1$, $n+2$ и т. д. Уничтожая последовательно все высшие гомотопические группы, мы и получаем некоторый комплекс B , для которого $\pi_i(B) = 0$ при $i \neq n-1$ и $\pi_{n-1}(B) = \mathbb{Z}$. Хорошо известно, что любые два полиэдра, являющиеся пространствами типа $K(\pi, n-1)$, при заданных $n-1$ и π гомотопически эквивалентны. Ясно, что $H^{n-1}(K(\mathbb{Z}, n-1), \mathbb{Z}) = \mathbb{Z}$, $K(\mathbb{Z}, n-1) = \sigma^0 \cup \sigma^{n-1} \cup \sigma^{n+1} \cup \sigma^{n+2} \cup \dots$. Пусть $\alpha \in H^{n-1}(K(\mathbb{Z}, n-1), \mathbb{Z})$ — образующая группы $\mathbb{Z}$. Тогда (см. [4]) функция, относящая отображению $f: K \rightarrow K(\mathbb{Z}, n-1)$ класс когомологий $f^*\alpha \in H^{n-1}(K, \mathbb{Z})$, для любого клеточного комплекса K определяет взаимно-однозначное соответствие между $H^{n-1}(K, \mathbb{Z})$ и множеством $\Pi(K, K(\pi, n))$ гомотопических классов отображений K в $K(\pi, n)$.

Условие теоремы 3, согласно которому гомоморфизм $i_*: H_{n-1}(S^{n-1}, U) \rightarrow H_{n-1}(K, U)$ не имеет ядра, эквивалентным образом может быть переформулировано на языке когомологий, а именно гомоморфизм $i^*: H^{n-1}(K, \mathbb{Z}) \rightarrow H^{n-1}(S^{n-1}, \mathbb{Z})$, индуцированный вложением, является эпиморфизмом. Следовательно, существует элемент $\beta \in H^{n-1}(K, \mathbb{Z})$ такой, что $i^*\beta = \alpha \in H^{n-1}(S^{n-1}, \mathbb{Z})$. В силу предыдущего существует непрерывное отображение $f: K \rightarrow K(\mathbb{Z}, n-1)$ такое, что $f^*\alpha = \beta$. Так как $(fi)^*\alpha = \alpha$, то ограничение отображения f на сферу iS^{n-1} , вложенную в K , гомотопно тождественному отображению сферы S^{n-1} на себя. Осталось доказать, что отображение f можно продеформировать в такое отображение, которое останется тождественным на сфере S^{n-1} и будет отображать на нее весь комплекс K . Но так как комплекс K является, по предположению, n -мерным и так как комплекс $K(\mathbb{Z}, n-1)$ не содержит клеток размерности n , то из теоремы о клеточной аппроксимации [4] можно считать, что отображение f переводит K в $(n-1)$ -мерный остов пространства $K(\mathbb{Z}, n-1)$, т. е. в сферу S^{n-1} . Этим и завершается доказательство теоремы 3.

§ 3. Расслоения

1. Определение локально тривиального расслоения.

Во многих прикладных задачах после их переформулировки на геометрическом языке возникают пространства E , снабженные специальной структурой, основное свойство которой можно

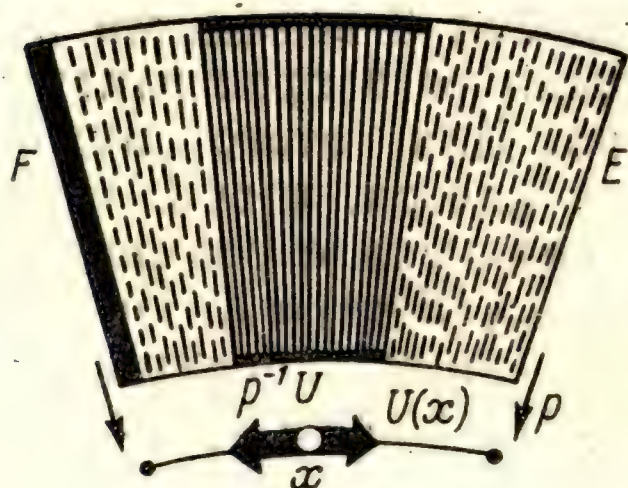


Рис. 22

грубо сформулировать так: пространство E разбивается в непрерывное объединение своих подпространств (слоев), гомеоморфных одному и тому же пространству F , причем это разбиение устроено локально довольно просто, а именно любое подмножество точек в E , составленное из «ближних слоев», гомеоморфно прямому произведению слоя F на некоторое пространство (рис. 22). Такое расслоение E на слои F оказывается чрезвычайно

полезным при изучении геометрии пространства E . Дадим теперь математическое определение.

Определение 1. Пусть нам даны топологические пространства E, B, F и непрерывное отображение p пространства E на B . Будем говорить, что четверка (E, B, F, p) является локально тривиальным расслоением, если для каждой точки $b \in B$ существуют открытая окрестность $U \subset B$ и гомеоморфизм $\varphi: p^{-1}U \rightarrow U \times F$ такие, что следующая диаграмма коммутативна:

$$\begin{array}{ccc} p^{-1}U & \xrightarrow{\varphi} & U \times F \\ \downarrow p & & \downarrow \pi \\ U & \xrightarrow{\pi} & U \end{array}$$

где $\pi: U \times F \rightarrow U$ — естественная проекция прямого произведения на сомножитель, коммутативна, т. е. $\pi\varphi = p$. Через $p^{-1}U$ обозначен полный прообраз U при отображении p . При этом E называется пространством расслоения (расслоенным пространством), F — слоем, B — базой, p — проекцией.

Иногда называют расслоением само отображение p , используя обозначение $p: E \rightarrow B$.

Определение 2. Расслоение $p: E \rightarrow B$ называется тривиальным, если E гомеоморфно $B \times F$, причем диаграмма коммутативна, т. е. $p = \pi\alpha$, где $\alpha: E \rightarrow B \times F$ — гомеоморфизм, $\pi: B \times F \rightarrow B$ — проекция.

Последнее требование означает, что при гомеоморфизме α каждый слой соответствует одной точке базы (после проекции). Происхождение термина «локальная тривиальность» видно из определения 2: расслоение $p^{-1}U \rightarrow U$ тривиально для достаточно малой окрестности U .

2. Примеры расслоений.

Пример 1. Прямые произведения $E = B \times F$ являются расслоениями относительно обеих проекций $B \times F \rightarrow B$ и $B \times F \rightarrow F$ (рис. 23).

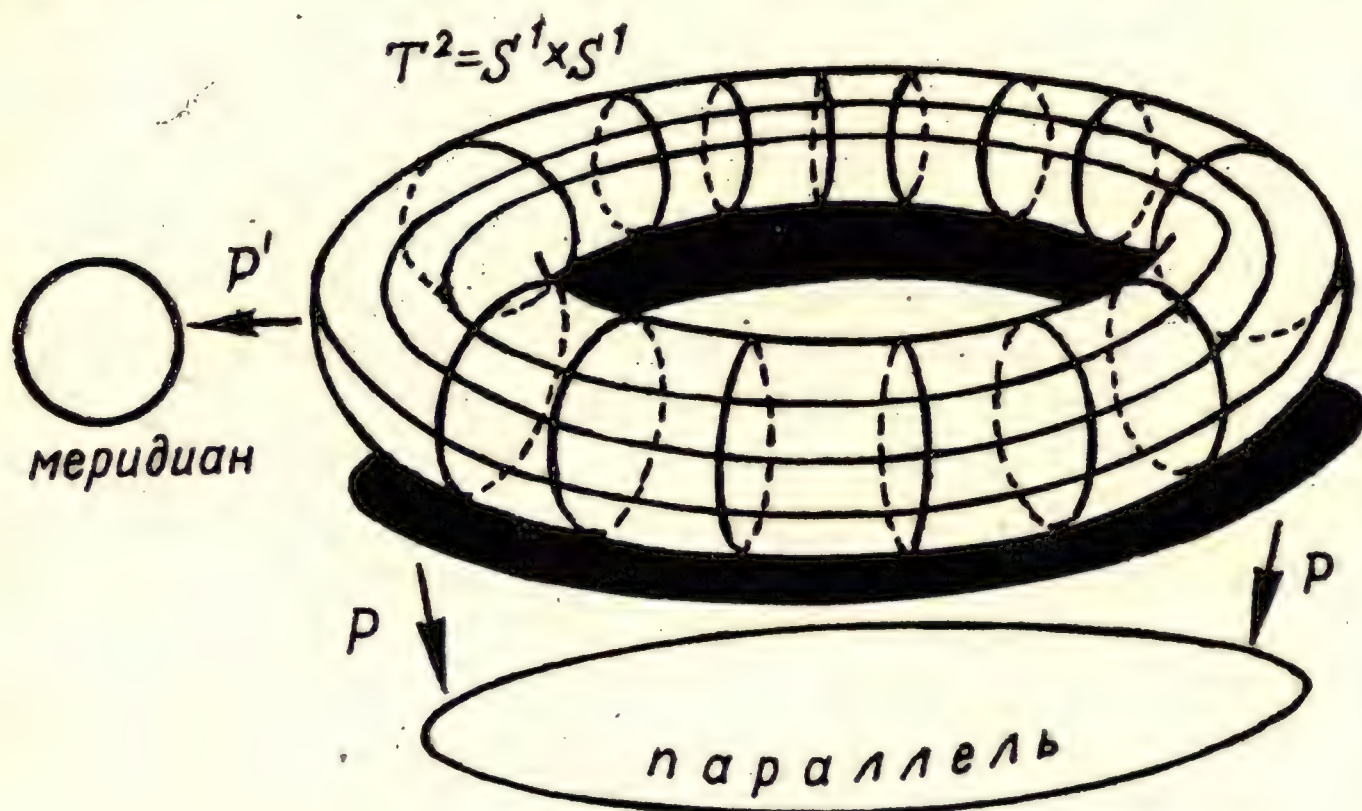


Рис. 23

Пример 2. Лист Мебиуса E , где F — отрезок I , а B — окружность S^1 , $p: E \rightarrow B$ — проекция листа Мебиуса на его среднюю линию (рис. 24). Так как E — неориентируемое много-

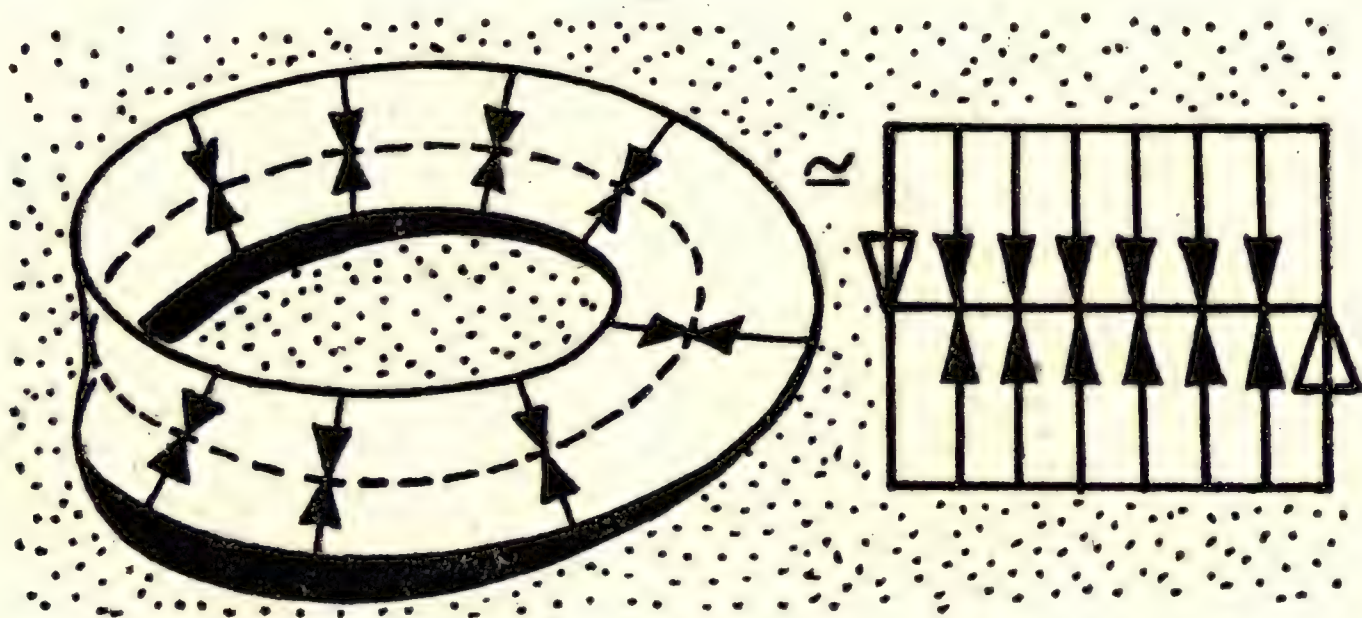


Рис. 24

образие, то это расслоение не является тривиальным (докажите!).

Пример 3. Пусть $f: M^k \rightarrow N^p$ — гладкое отображение связного компактного гладкого многообразия M на гладкое компактное многообразие N , причем $k \geq p$ и f — регулярное отображение в том смысле, что его дифференциал df имеет максимальный ранг, равный p в каждой точке $x \in M$. Тогда, как известно (см., например, [1]), каждый слой $F_y = f^{-1}(y)$, $y \in N$ является гладким $(k-p)$ -мерным подмногообразием в M , и, сле-

довательно, учитывая теорему о неявных функциях, получаем, что отображение f определяет локально тривиальное расслоение.

Этот пример особенно важен, так как большинство расслоений, с которыми мы будем встречаться, именно таковы. Среди них выделен важный класс, связанный с однородными пространствами. Пусть G — компактная группа Ли и $H \subset G$ — замкнутая подгруппа [1]. Рассмотрим множество G/H классов смежности gH по подгруппе H . Это множество естественно снабжается топологией (близкие классы смежности определяются как сдвиги подгруппы H на близкие элементы $g, g' \in G$). Получаем отображение $p: G \rightarrow G/H$, где $p(g) = (gH) \in G/H$. Пространство G/H естественно снабжается структурой гладкого многообразия [1]. Мы не будем здесь это доказывать, так как в дальнейших примерах этот факт очевиден.

3. Геометрия расслоения Хопфа.

Пример 4. Расслоение Хопфа. Это классическое расслоение встречается во многих приложениях и является частным случаем примера 3. Рассмотрим сферу S^3 , стандартно вложенную в комплексное пространство $\mathbb{C}^2(z_1, z_2)$ как множество то-

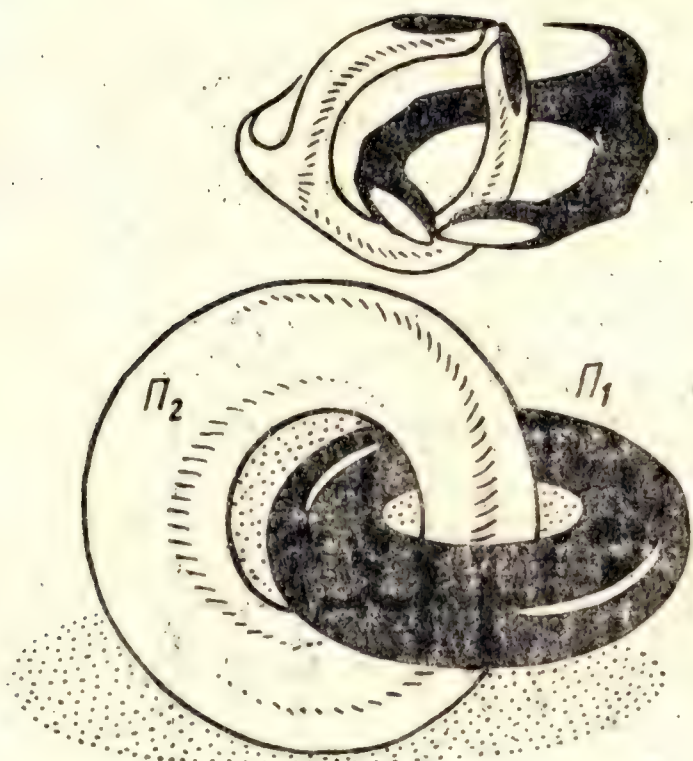


Рис. 25

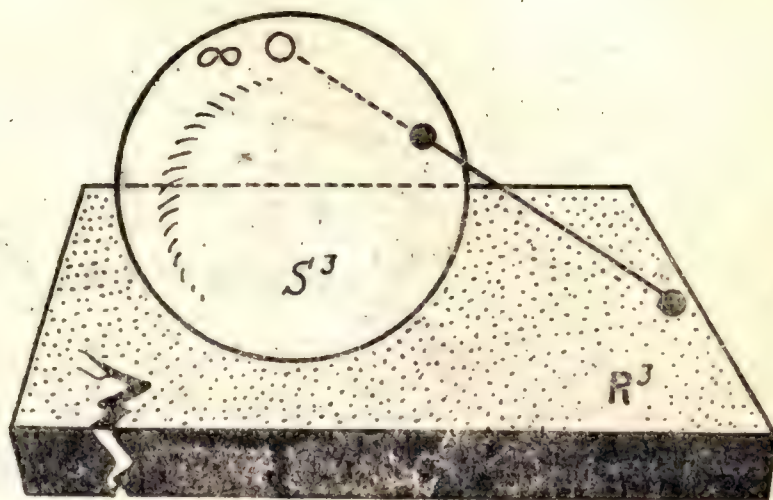


Рис. 26

чек (z_1, z_2) таких, что $|z_1|^2 + |z_2|^2 = 1$. Рассмотрим двумерную сферу S^2 как плоскость $\mathbb{R}^2$, пополненную бесконечно удаленной точкой, и отождествим S^2 с комплексной проективной прямой $\mathbb{CP}^1$. Построим отображение $p: S^3 \rightarrow S^2$, задаваемое так: $p(z_1, z_2) = z_1/z_2$. Легко видеть, что p — гладкое отображение и что оно регулярно во всех точках. Следовательно, p определяет локально тривиальное расслоение с базой S^2 и слоем S^1 , так как $p(e^{i\varphi}z_1, e^{i\varphi}z_2) = p(z_1, z_2)$. Расслоение Хопфа можно описать наглядно, используя представление сферы S^3 в виде объединения двух полноторий. В самом деле, рассмотрим в $\mathbb{C}^2$ конус K , задаваемый уравнением $|z_1| = |z_2|$. Его пересечение со сферой S^3 гомеоморфно тору $T^2 = S^1 \times S^1$. Следовательно,

сфера разбивается в сумму замкнутых подмножеств Π_1 и Π_2 , где $\Pi_1 = \{|z_1| \leq |z_2|\}$, $\Pi_2 = \{|z_1| \geq |z_2|\}$. Они пересекаются по двумерному тору. Каждое из множеств Π_i гомеоморфно полноторию $S^1 \times D^2$ (рис. 25). Рисунок условен в том смысле, что по-

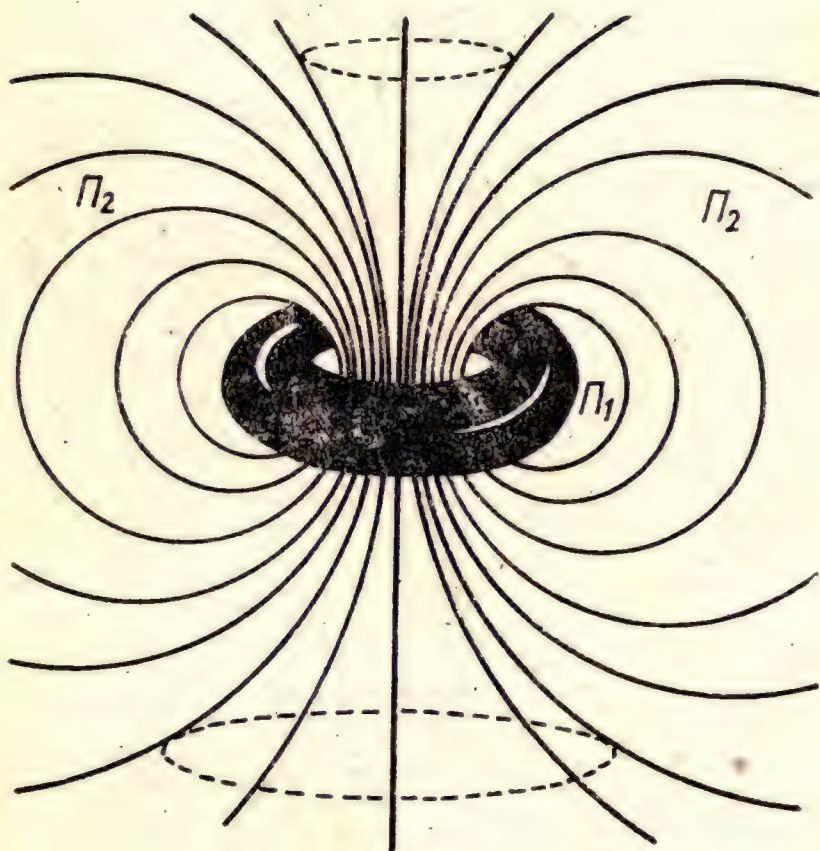


Рис. 27

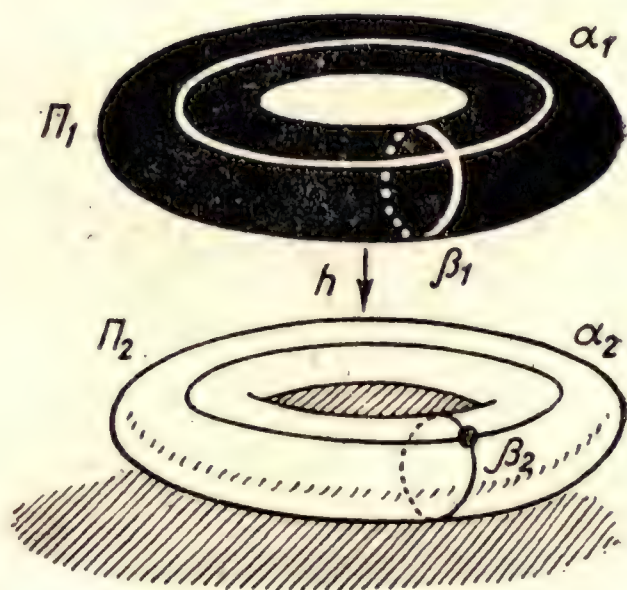


Рис. 28

казывает полнотория в $\mathbf{R}^3$, а не в S^3 . Чтобы изображение стало точным, следует представить S^3 как $\mathbf{R}^3$, пополненное бесконечно удаленной точкой (рис. 26). В этом представлении можно считать полноторие Π_1 стандартно вложенным в $\mathbf{R}^3$ (рис. 27), тогда дополнение к нему в $S^3 = \mathbf{R}^3 \cup \infty$ гомеоморфно второму полноторию Π_2 . Если фиксировать на полноториях стандартные параллели и меридианы α_1, β_1 и α_2, β_2 (рис. 28), то склейка полноторий по их границе задается диффеоморфизмом $h: T^2 \rightarrow T^2$, переводящим параллели в меридианы и наоборот. Индуцированное отображение одномерных групп гомологий $H_1 = \mathbf{Z} \oplus \mathbf{Z}$ тора задается целочисленной матрицей $h_* =$

$$= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad h(\alpha_1) = -\beta_2, \quad h(\beta_1) = \alpha_2.$$

Теперь можно наглядно описать расслоение Хопфа. Для этого достаточно предъявить расслоение сферы S^3 на окружности. Расслоим $\mathbf{R}^3 \cup \infty = S^3$ на семейство концентрических торов (рис. 29). Для наглядности мы рассекли $\mathbf{R}^3$ двумерной плоскостью (z, x) . При этом возникают две предельные ситуации. Первая: концентрические торы сжимаются на стандартную окружность, лежащую в плоскости (y, x) . Обозначим ее S_1^1 ; она пунктиром показана на рис. 29. Вторая предельная ситуация: торы, расширяясь, стремятся к вертикальной оси Oz , являющейся их «пределом», так как вся бесконечность склеена в одну точку. Ось Oz изображает окружность S_2^1 в сфере S^3 . Ясно, что S_1^1 — ось полнотория Π_1 , а S_2^1 — ось полнотория Π_2 . Для описания слоев расслоения

Хопфа фиксируем какой-либо из торов, расслаивающих $\mathbb{R}^3 \cup \infty$. Зададим на торе гладкую окружность, обходящую его один раз вокруг параллели и один раз вокруг меридиана. Поворачивая затем эту траекторию ортогональным преобразованием, сохраняющим ось Oz , мы расслаиваем тор в объединение непересекающихся окружностей (рис. 30). Это же расслоение показано

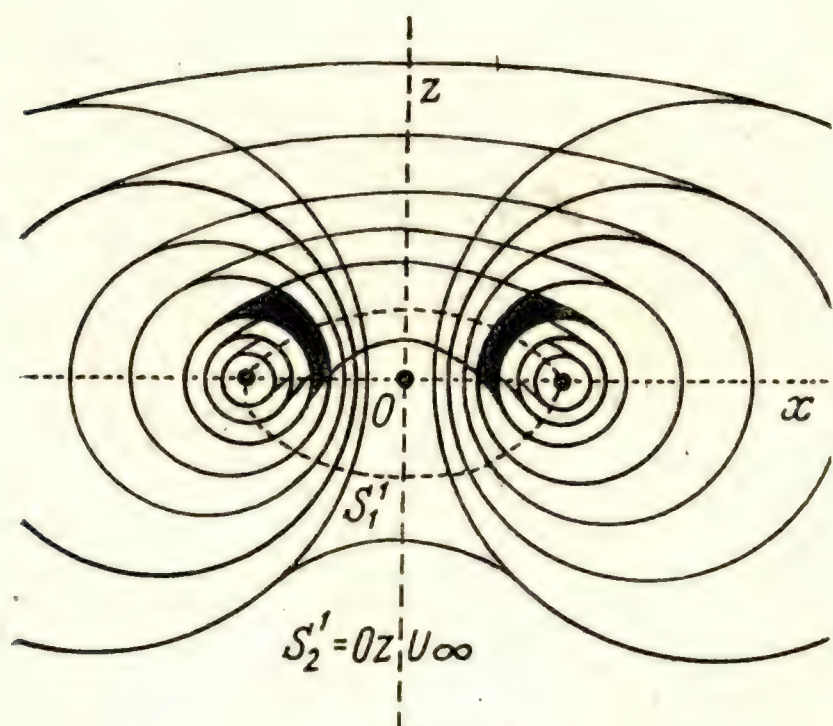


Рис. 29

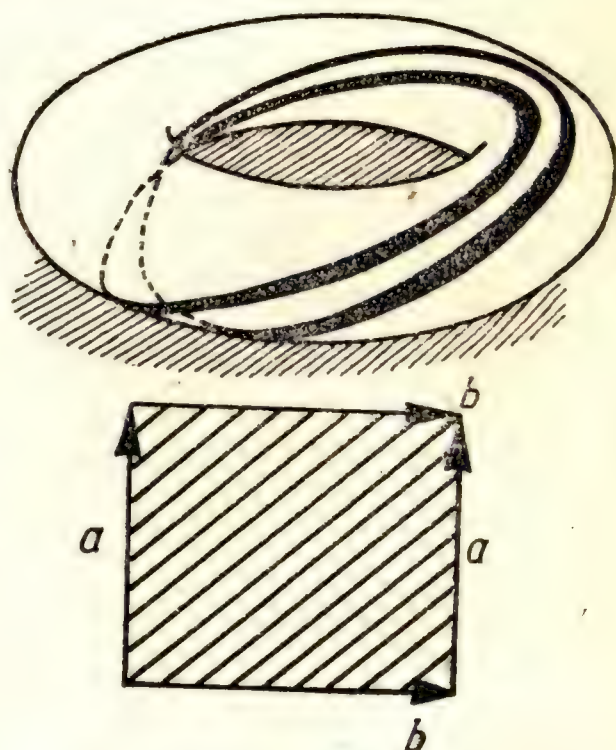


Рис. 30

на рис. 30 на модели тора — на квадрате с отождествлениями на границе. Такое же расслоение построим на каждом из торов, объединение которых дает трехмерную сферу. Так как они не пересекаются, то расслаивающие их окружности также не пересекаются. В обоих предельных случаях, когда торы стремятся к S_1^1 и к S_2^1 , расслаивающие их окружности стремятся к S_1^1 и S_2^1 соответственно. Это и есть расслоение S^3 , задающее расслоение Хопфа (проверьте!). Любые два слоя этого расслоения зацеплены, т. е. двумерная пленка, например диск, натянутая на один из слоев, обязательно пересекается с другим слоем. Это обстоятельство хорошо видно на рис. 30. См. также рис. 31(а). Из построенного нами изображения расслоения Хопфа видно, в какие точки базы, двумерной сферы S^2 переходят те или иные слой-окружности. Рассмотрим диск D^2 в плоскости (x, y) с границей S_1^1 . Ясно, что любая окружность-слой, отличная от S_1^1 , обязательно пересекает диск ровно в одной точке (рис. 31(б)). Когда слой стремится к слою S_1^1 , точка его пересечения с диском стремится к границе диска, и когда слой совпадает со слоем S_1^1 , мы должны отождествить окружность $S_1^1 = \partial D^2$ в одну точку. При этом диск D^2 превращается в двумерную сферу, каждой точке которой отвечает одна и только одна окружность-слой расслоения Хопфа. Слой S_2^1 соответствует при этом центру диска (рис. 31). Итак, для построения расслоения Хопфа нужно расслоить S^3 на окружности описанным способом, а затем сопоставить каждой из них точку ее пересечения с диском D^2 (или сразу всю границу диска, если

$S^1 = S_1^1$). Расслоение Хопфа не является тривиальным, в противном случае (т. е. если бы $S^3 = S^1 \times S^2$) группа $H_1(S^3, \mathbb{Z})$ была бы изоморфна группе $\mathbb{Z}$ (докажите!), что невозможно, поскольку $H_1(S^3, \mathbb{Z}) = 0$.

Мы остановились на этом примере расслоения столь подробно, так как здесь видны многие важные эффекты, характери-

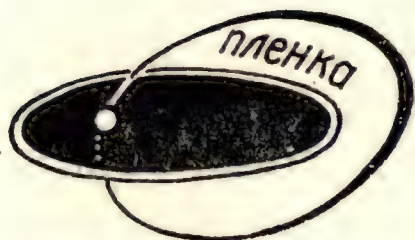


Рис. 31(а)

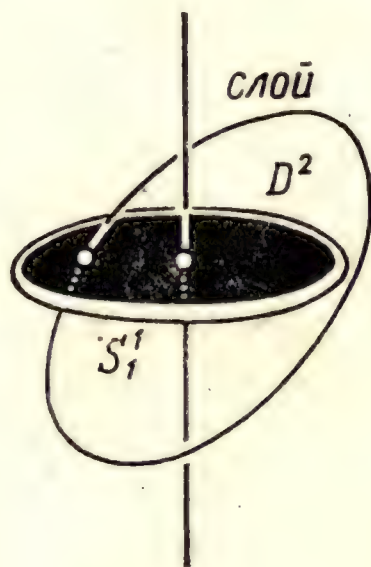


Рис. 31(б)

зующие расслоения общего вида. Приведем еще один полезный пример расслоения.

4. Геометрия расслоения единичных касательных векторов к сфере и индекс векторного поля.

Пример 5. Пусть E — пространство всех единичных касательных векторов к четномерной сфере S^{2n} , база B — это сфера S^{2n} , проекция $p: E \rightarrow B$ сопоставляет касательному вектору

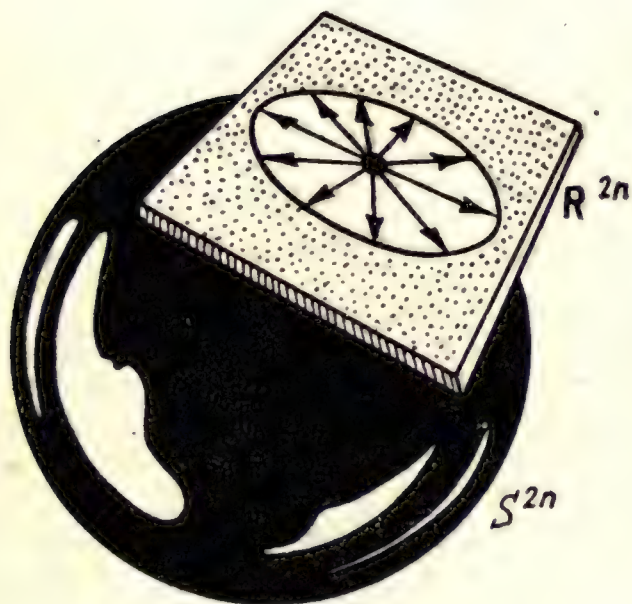


Рис. 32(а)

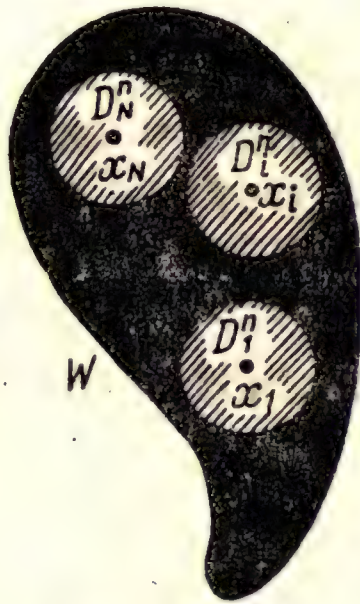


Рис. 32(б)



Рис. 32(в).

его начальную точку, т. е. точку касания к сфере. Это локально тривиальное расслоение со слоем S^{2n-1} (рис. 32(а)).

Лемма 1. Расслоение $p: E \rightarrow S^{2n}$ не является тривиальным.

Доказательство. Допустим противное, а именно, что $S^{2n} \times S^{2n-1} = E$. Коммутативность диаграммы из определения 1 означает, что существует ненулевое сечение этого расслоения, т. е. непрерывное отображение $h: S^{2n} \rightarrow E$ такое, что $ph = 1_{S^{2n}}$ (тождественное отображение базы). Но это означало бы, что в каждой точке $b \in S^{2n}$ был бы однозначно определен ненулевой (единичной длины) касательный вектор $h(b)$, являющийся непрерывной функцией точки b . Осталось доказать, что на S^{2n} нет непрерывного касательного векторного поля, отличного от нуля в каждой точке. Это вытекает из общей теоремы об индексе векторного поля на сфере, имеющей и другие приложения.

Напомним определение индекса векторного поля. Пусть $v(x)$ — гладкое поле на гладком многообразии M^n и пусть поле имеет лишь конечное число особых точек x_0 , т. е. таких, где поле равно нулю. Тогда эти точки изолированы, т. е. для каждой из них существует открытая окрестность, в которой нет других нулей поля. Рассмотрим в M^n сферу S_ε^{n-1} достаточно малого радиуса ε с центром в особой точке x_0 . Если ε мало, то можно считать, что $v(x) \neq 0$ на всей сфере S_ε . Переходя к локальным координатам $x_1, \dots, x_n$ в окрестности точки x_0 , можно считать, что сфера S_ε лежит в области в $\mathbb{R}^n$, следовательно, определено отображение $f: S_\varepsilon \rightarrow S$, где $f(x) = \frac{v(x)}{|v(x)|}$, $|v(x)|$ — евклидова длина $v(x)$.

Определение 3. Индексом $\text{ind}_{x_0} v$ особой точки x_0 векторного поля v называется целое число $\deg f$, т. е. степень отображения $f: S_\varepsilon \rightarrow S$. Сумма индексов особых точек поля по всем его особым точкам называется индексом векторного поля $\text{ind } v$.

Известно [1], что это определение корректно и не зависит от выбора локальных координат в окрестности особой точки.

Теорема 1. Пусть $W^n \subset \mathbb{R}^n$ — ограниченная область в $\mathbb{R}^n$, причем граница ∂W области W является гладким компактным замкнутым подмногообразием размерности $n-1$ в $\mathbb{R}^n$. Пусть v — гладкое векторное поле на W , причем $v(x) \neq 0$ на границе ∂W и v имеет лишь конечное число особых точек в области W . Тогда индекс этого поля равен $\deg F$, где $F: \partial W \rightarrow S^{n-1}$ гладкое отображение, задаваемое так: $F(x) = \frac{v(x)}{|v(x)|}$, где $x \in \partial W$.

Следствие 1. Если векторные поля v_1 и v_2 на W таковы, что $\deg F = \deg F_{v_2}$ на ∂W , то индексы этих полей совпадают. В частности, если поле v_2 получено из поля v_1 гладкой гомотопией g_t такой, что $g_t v_1(x) \neq 0$ на границе ∂W при $0 \leq t \leq 1$, то индексы полей v_1 и v_2 совпадают.

Доказательство теоремы 1. Пусть $x_1, \dots, x_N$ — особые точки поля v , $N < \infty$. Окружим каждую особую точку x_i сферой S_i^{n-1} достаточно малого радиуса так, чтобы $D_i \cap D_j = \emptyset$ при $i \neq j$, где D_i — n -мерный шар, ограниченный сферой S_i и

имеющий центром точку x_i (рис. 32(б)). Рассмотрим область V , получающуюся из W выбрасыванием всех шаров D_i (рис. 32(в)). Построим гладкое отображение $f: V \rightarrow S$, положив $f(x) = \frac{v(x)}{|v(x)|}$, где S — стандартная сфера единичного радиуса в $\mathbb{R}^n$. Это отображение корректно определено, так как поле v отлично от нуля во всех точках области V . Рассмотрим на сфере S стандартную форму риманова объема ω , инвариантную относительно ортогональных преобразований $\mathbb{R}^n$. Тогда интеграл $\int_S \omega$ равен объему сферы S , который мы обозначим

через a_{n-1} . Форма ω , являясь внешней дифференциальной формой на сфере, индуцирует при отображении f некоторую форму $f^*\omega$ на V . Рассмотрим интеграл $\int_V df^*\omega$, где d — операция внешнего дифференцирования [5, гл. 6]. По формуле Стокса:

$\int_V df^*\omega = \int_{\partial V} f^*\omega$, где ∂V — граница области V (знак учтен выбором ориентации). Но граница ∂V распадается в несвязное объединение следующих $(n-1)$ -мерных подмногообразий $\partial W \cup (-\cup S_i)$. Знак «—» указывает ориентацию, противоположную индуцированной. Мы получаем:

$\int_V df^*\omega = \int_{\partial W} f^*\omega - \sum_i \int_{S_i} f^*\omega$.

Из свойств оператора d [5, гл. 6] следует, что он перестановочен с операцией индуцирования форм, т. е. $df^* = f^*d$. Итак, $\int_V df^*\omega = \int_V f^*d\omega$.

Но форма ω имеет максимальную размерность на сфере S , т. е. $d\omega = 0$, т. е. $\int_V df^*\omega = 0$. Итак, $\int_{\partial W} f^*\omega = \sum_i \int_{S_i} f^*\omega$. По теореме 2 из

[5, с. 394] мы имеем: $\int_{\partial W} f^*\omega = \deg f|_{\partial W} \cdot \int_S \omega = a_{n-1} \cdot \deg f|_{\partial W}$. Точно

так же получаем: $\int_{S_i} f^*\omega = a_{n-1} \cdot \deg f|_{S_i}$. Итак, $a_{n-1} \cdot \deg f|_{\partial W} = a_{n-1} \times$

$\times \sum_i \deg f|_{S_i}$. Из определения индекса векторного поля получаем:

$\deg F = \sum_i \text{ind}_{x_i} v = \text{ind } v$. Теорема доказана.

Теорема 2. Любое непрерывное касательное векторное поле v на сфере S^{2n} обязательно имеет хотя бы одну особую точку x_0 , т. е. такую, что $v(x_0) = 0$.

Доказательство. Допустим противное, т. е. что на сфере S^{2n} существует касательное поле, отличное от нуля во всех точках. В силу теоремы 1 из [1, с. 478] это поле сколь угодно близко аппроксимируется гладким касательным полем, которое обозначим тем же символом v . Так как исходное поле было

ненулевым, то и аппроксимирующее его гладкое поле также отлично от нуля во всех точках. Теорему 2 мы получим как следствие следующего более общего утверждения.

Теорема 3. Пусть v — гладкое касательное векторное поле на сфере S^k , имеющее, быть может, лишь изолированные особые точки. Тогда его индекс равен $1 + (-1)^k$, т. е. равен нулю на нечетномерной сфере и двум на четномерной сфере.

Отсюда следует теорема 2, так как выше мы, предположив противное утверждению теоремы, построили на S^{2n} касательное поле без особенностей, т. е. поле нулевого индекса, что невозможно по теореме 3.

Доказательство теоремы 3. Пусть $x_0 \in S^k$ — произвольная точка и D_1 — диск в сфере S^k достаточно малого радиуса, имеющий точку x_0 своим центром (рис. 33(a)). Предста-

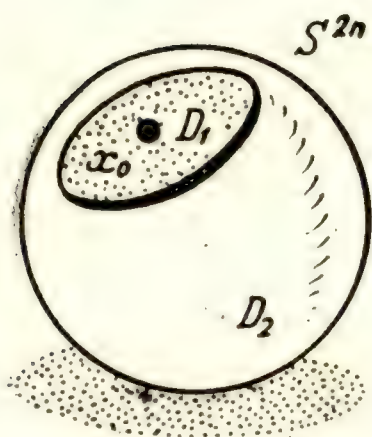


Рис. 33(а)

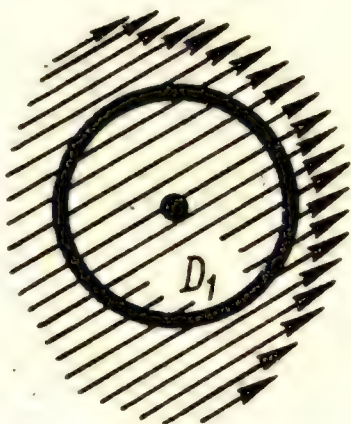


Рис. 33(б)

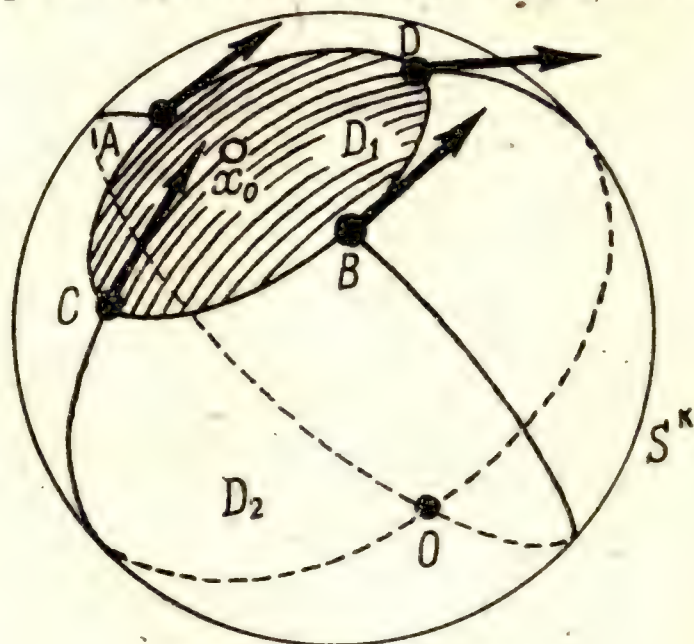
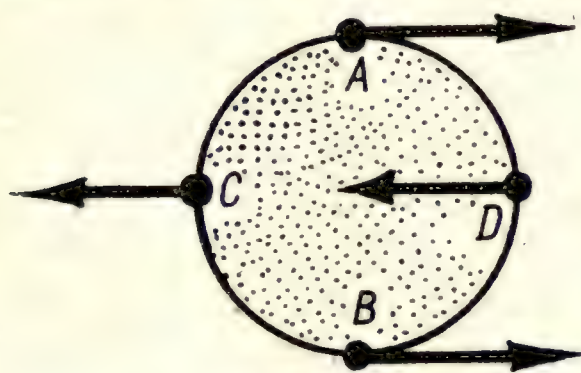


Рис. 33(в)

вим S^k в виде объединения двух замкнутых дисков $D_1 \cup D_2$, где D_2 — замыкание дополнения к D_1 в S^k . Пусть x_0 — неособая точка поля v , т. е. $v(x_0) \neq 0$. Можно считать, что на всем диске D_1 поле отлично от нуля и, более того, сколь угодно близко к «параллельному полю», т. е. интегральные траектории поля v сколь угодно близки к пучку параллельных прямых (рис. 33). Так как поле не имеет особых точек на D_1 , то все особые точки (если они есть) сосредоточены на диске D_2 . Этот диск можно перевести при помощи разворачивающего диффеоморфизма (например, стандартной стереографической проекцией с центром в точке x_0) на стандартный диск в $\mathbb{R}^k$ с центром в точке O . На рис. 33 эта операция эквивалентна следующей: нужно выбросить из сферы диск D_1 и оставшийся диск развернуть и поместить на плоскость. Проследим за тем, как изобразятся векторы поля v на границе диска D_2 после этой операции, т. е. после введения на диске D_2 евклидовых координат. Ясно, что с точки этих координат в диске D_2 поле на его границе примет вид, показанный на рис. 34(а). Уже сейчас интуитивно ясно, что, например, в двумерном случае наличие такого завихрения поля на границе диска означает, что внутри него есть по крайней мере одна особенность. Сейчас мы установим этот факт на

основе теоремы 1. Диск $D_2 \subset \mathbb{R}^k$ можно отождествить с областью W из теоремы 1. Тогда граница ∂W — это сфера S^{k-1} , а поле v на ∂W устроено так, как показано на рис. 34(а) (для двумерного случая). Из теоремы 1 следует, что для нахождения индекса поля на диске D_2 достаточно вычислить степень отображения $F: S^{k-1} \rightarrow S^{k-1}$, где $F(x) = v(x)/|v(x)|$. Для этого достаточно рассмотреть регулярную точку и вычислить знаки якобианов отображения во всех ее прообразах. Так как степень



отображения не зависит от выбора регулярной точки, то для удобства можем взять точку C на сфере S^{k-1} . При этом мы отождествляем два экземпляра сферы S^{k-1} и считаем, что отображение F переводит сферу в себя (рис. 34(а)). Из построения поля v следует,

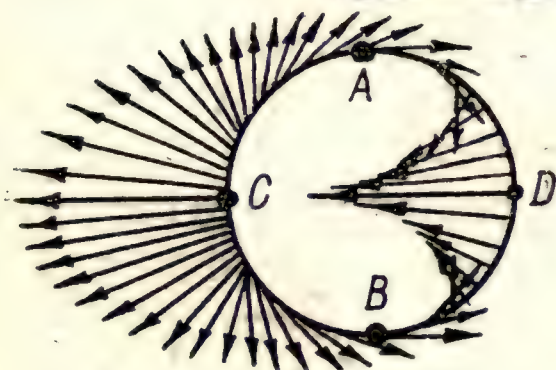


Рис. 34(а)

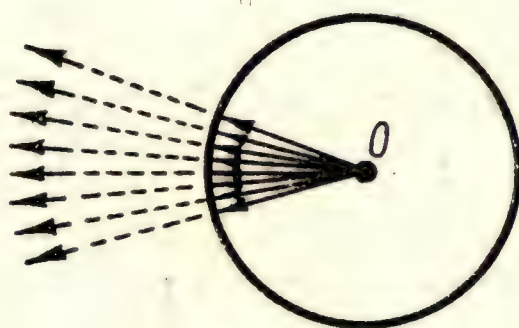


Рис. 34(б)

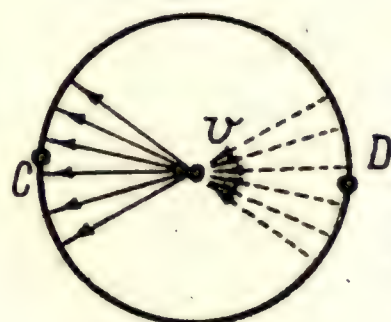


Рис. 34(в)

что прообразами точки C являются только две точки: сама точка C и диаметрально противоположная ей точка D . Отображение F в точке C показано на рис. 34(б). Его дифференциал, определенный на касательной плоскости к сфере в точке C , задается диагональной матрицей с положительными числами на диагонали. Следовательно, знак якобиана в точке C положителен. В точке D ситуация иная. Каждый вектор v , близкий к точке D , отображается в вектор v' , выходящий из точки O и получающийся из v параллельным переносом в точку O (рис. 34(в)). Поэтому дифференциал dF в точке D записывается на касательной плоскости к сфере в виде диагональной матрицы, по диагонали которой стоят отрицательные числа. Итак, мы имеем дело с отображением сферы относительно ее центра. Степень отображения $\alpha(x) = -x$ сферы S^{k-1} на себя равна $(-1)^k$. Итак, знак якобиана в точке D равен $(-1)^k$, что и доказывает теорему 3.

В научно-популярной литературе теорему 3 иногда называют «теоремой о еже»: свернувшегося в клубок ежа нельзя причесать так, чтобы все его иголки легли по касательной (если k четно).

Итак, касательное расслоение $p: E \rightarrow S^{2n}$ не является тривиальным. Расслоение $p: E \rightarrow S^1$ тривиально, так как гомеоморфно $S^1 \times S^0$, где S^0 — нульмерная сфера, пара точек.

Выше мы подробно изучили некоторые свойства расслоения Хопфа, поскольку оно потребуется нам в дальнейших главах. Укажем еще одно любопытное свойство этого расслоения. Если гомотопические группы конечного клеточного комплекса тривиальны, то он стягивается в точку (если он связан). Если комплекс односвязен, то этот же факт следует из тривиальности групп целочисленных гомологий. Вопрос: будет ли непрерывное отображение одного комплекса в другой гомотопно отображению в точку, если оно индуцирует нулевой гомоморфизм всех гомотопических групп (или всех групп гомологий)? Оказывается, нет. Рассмотрим расслоение Хопфа $p: S^3 \rightarrow S^2$. Пусть $T^3 = S^1 \times S^1 \times S^1$ — трехмерный тор. Стягивая в точку его двумерный остов, мы получаем непрерывное отображение $g: T^3 \rightarrow S^3$. Беря композицию $f = gp$, мы получаем отображение $f: T^3 \rightarrow S^2$, которое, как легко проверить, не гомотопно отображению в точку. В самом деле, если бы отображение f было гомотопно нулю, то по теореме о накрывающей гомотопии эту гомотопию можно было бы накрыть некоторой гомотопией в S^3 . В результате мы получили бы, что отображение g гомотопно отображению, которое переводит тор в слой расслоения Хопфа, т. е. в окружность. Но так как эта окружность стягивается по сфере S^3 в точку, то отображение g также оказалось бы гомотопным нулю, что противоречит его построению. В то же время легко видеть, что индуцированные отображения целочисленных гомологий и когомологий тривиальны. Более того, тривиальны и индуцированные отображения гомотопических групп $f_*: \pi_i(T^3) \rightarrow \pi_i(S^2)$, поскольку двумерная сфера односвязна, а гомотопические группы тора T^3 равны нулю, начиная с размерности два и выше. Аналогичный пример можно построить и в том случае, когда оба комплекса односвязны.



§ 1. Точные функции Морса

Настоящий параграф в некоторой степени опирается на материал, изложенный в [1, 3, 34], и является непосредственным продолжением раздела в [34], посвященного изучению функций Морса на многообразии.

Вкратце напомним, что функция Морса характеризуется тем, что все ее критические точки невырожденные. Критическая точка x функции f называется невырожденной, если в окрестности точки x существует регулярная система координат, в которой функция f имеет вид диагональной квадратичной формы от всех переменных. Число минусов квадратичной диагональной формы называется индексом критической точки. Функция Морса общего положения — это такая функция Морса, которая принимает различные значения в критических точках. В дальнейшем мы будем в основном рассматривать функции Морса общего положения. Напомним также, что функция Морса называется правильной, если для любых двух критических точек x индекса λ и y индекса $\lambda+1$ выполнено строгое неравенство $f(x) < f(y)$. Обозначим через F_λ множество всех функций Морса, у которых минимальное (для данного многообразия) число критических точек индекса λ .

О п р е д е л е н и е 1. Точной или минимальной функцией Морса на многообразии M^n называется функция, имеющая минимальное число критических точек в каждой размерности, т. е.

$$f \in \bigcap_{\lambda=0}^n F_\lambda.$$

Другими словами, для всех λ число критических точек индекса λ у точной функции Морса является наименьшим возможным в классе всех функций Морса на данном многообразии. Если $N_\lambda(f)$ — число критических точек индекса λ , то для точной функции Морса f имеет место равенство $N_\lambda(f) = \min N_\lambda(g)$, где минимум берется по всем функциям Морса g на многообразии. Вообще говоря, точной функции Морса на данном многообразии может не существовать. Оказывается, если многообразие неодносвязно, то существуют топологические инварианты, в некоторых случаях препятствующие существованию на многообразии M хотя бы одной точной функции Морса (в смысле определения 1). Однако для односвязных многообразий достаточно большой размерности точная функция Морса всегда существует (см. ниже).

Вопрос о существовании точной функции Морса естественно возникает при анализе топологической структуры данного многообразия. Напомним, что с каждой функцией Морса на M естественно связано некоторое разбиение многообразия M в объединение ручек [2, 3, 34]. В свою очередь такое представление позволяет описать гомотопическую структуру многообразия, представив его в виде клеточного комплекса, в котором число клеток размерности λ равно числу критических точек индекса λ у функции Морса. Разным функциям Морса отвечают, вообще говоря, различные разложения одного и того же многообразия в сумму ручек. Минимизируя число критических точек, мы тем самым минимизируем количество ручек, т. е. получаем более простое представление многообразия в виде «суммы элементарных кирпичей». Следовательно, задача нахождения функции Морса с минимальным числом критических точек возникает из задачи построения клеточного разбиения многообразия в сумму минимального числа клеток.

Рассмотрим представление абелевой группы гомологий $H_\lambda(M^n, \mathbb{Z})$ в виде прямой суммы свободной абелевой группы G_λ и конечной группы K_λ . Обозначим через p^λ ранг группы G_λ , т. е. минимальное количество образующих в группе G_λ , а через q^λ — минимальное число образующих в группе K_λ .

Теорема 1 (см. [51]). Пусть M^n — односвязное гладкое замкнутое многообразие размерности $n \geq 6$. Тогда на M^n существует точная функция Морса. При этом число N_λ критических точек индекса λ вычисляется следующим образом: $N_\lambda = p^\lambda + q^\lambda + q^{\lambda-1}$.

Доказательство. Мы приведем здесь лишь краткую схему доказательства, отсылая заинтересованного читателя за деталями к [51] или [53]. Пусть $f: M^n \rightarrow I$ — правильная функция Морса на многообразии M , где I — единичный отрезок. Доказательство существования правильных функций см. в [3]. Пусть c_λ — такие числа, что $f(x) < c_\lambda < f(y)$ для любых критических точек x и y индексов λ и $\lambda+1$ соответственно. Поверхность уровня $\{f=c_\lambda\}$ отделяет друг от друга критические точки x и y . Через M_λ обозначим следующее многообразие (с краем): $M_\lambda = f^{-1}[0, c_\lambda]$ (рис. 1). Пусть $\xi(f)$ — градиентно-подобное векторное поле для функции f [3, 34]. Задание функции f и поля $\xi(f)$ определяет разложение многообразия в сумму ручек $D^\lambda \times D^{n-\lambda}$ [3, 34]. Получаем последовательность включений: $M_0 \subset M_1 \subset \dots \subset M_n = M^n$. Хорошо известно [3, 34], что

$$H_i(M_\lambda, M_{\lambda-1}, \mathbb{Z}) = \begin{cases} 0 & \text{при } i \neq \lambda, \\ \mathbb{Z} \oplus \dots \oplus \mathbb{Z} \text{ (} k \text{ слагаемых)} & \text{при } i = \lambda, \end{cases}$$

где k — число критических точек (ручек) индекса λ . Базис в группе $H_i(M_\lambda, M_{\lambda-1}, \mathbb{Z})$ задается ручками индекса λ с фиксированными ориентациями (рис. 2). Геометрически образующие

группы гомологий можно задать в виде сепаратрисных дисков D^λ размерности λ , состоящих из сепаратрис поля $\xi(f)$, входящих в критические точки индекса λ . Рассматривая эти диски по модулю подпространства $M_{\lambda-1}$, мы получаем относительные циклы, реализующие базис в группе относительных гомологий. Рассмотрим цепной комплекс $\{C, \partial\}$, составленный из групп $C_\lambda = H_\lambda(M_\lambda, M_{\lambda-1}, \mathbb{Z})$ и связывающих их гомоморфизмов ∂_λ , где ∂_λ является граничным оператором из точной последовательности следующей тройки $M_\lambda \supset M_{\lambda-1} \supset M_{\lambda-2}$, а именно

$$\partial_\lambda: H_\lambda(M_\lambda, M_{\lambda-1}, \mathbb{Z}) \rightarrow H_{\lambda-1}(M_{\lambda-1}, M_{\lambda-2}, \mathbb{Z}).$$

Лемма 1. Группы гомологий цепного комплекса $\{C, \partial\}$, т. е. группы $H_\lambda(C, \partial) = \text{Ker } \partial_\lambda / \text{Im } \partial_{\lambda+1}$, канонически изоморфны группам гомологий многообразия $H_\lambda(M, \mathbb{Z})$.

Доказательство см., например, в [3]. Эта лемма позволяет вычислять гомологии многообразия, исходя из информации о расположении ручек. Известно также (см. [3]), что оператор ∂_λ задается матрицей индексов пересечения правосторонних и левосторонних сфер, являющихся границами соответствующих сепаратрисных дисков индексов λ и $\lambda+1$.

Следующая лемма может быть названа леммой о взаимном уничтожении ручек

соседних индексов. Оказывается, в некоторых ситуациях ручки соседних индексов λ и $\lambda+1$ могут «съесть» друг друга, что эквивалентно взаимному уничтожению двух критических точек таких индексов. В результате получается функция Морса с меньшим количеством критических точек. В частности, эта операция позволяет уменьшать количество ручек в разложении многообразия в сумму ручек.

Лемма 2 (см. [54]). Пусть $M' = M_{\lambda-1} \cup_\phi D^\lambda \times D^{n-\lambda} \cup_\psi D^{\lambda+1} \times D^{n-\lambda-1}$, где $D^\lambda \times D^{n-\lambda}$ и $D^{\lambda+1} \times D^{n-\lambda-1}$ — ручки индексов λ и $\lambda+1$ соответственно, приклеенные к многообразию $M_{\lambda-1}$ по отображениям ϕ и ψ . Пусть $\pi_1(M^n) = 0$, $\lambda \geq 2$, $\lambda+1 \leq n-3$, $n \geq 6$. Предположим также, что индекс пересечения левосторонней сферы ручки $D^\lambda \times D^{n-\lambda}$ с правосторонней сферой ручки $D^{\lambda+1} \times D^{n-\lambda-1}$ равен ± 1 . Тогда существует диффеоморфизм $h: M \rightarrow M$, тождественный вне некоторой открытой окрестности многообразия M' в M , такой, что $h(M') = M_{\lambda-1}$. Другими словами, две указанные ручки взаимно уничтожились.

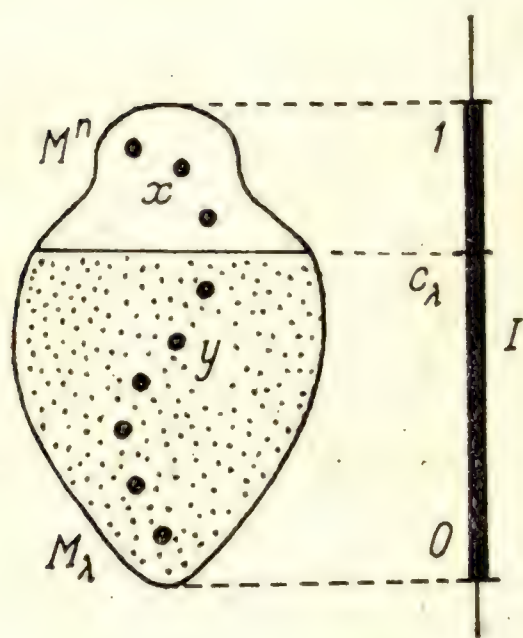


Рис. 1



Рис. 2

Напомним процесс сложения ручек. Рассмотрим многообразие $W' = M_{\lambda-1} \cup_{\varphi} D^{\lambda} \times D^{n-\lambda}$, получающееся из многообразия $M_{\lambda-1}$ приклеиванием ручки индекса λ . Предположим, что M^n односвязно и что $\lambda \geq 2$. Тогда отображение $\varphi|_{\partial D^{\lambda} \times 0}$ определяет в группе $\pi_{\lambda-1}(\partial M_{\lambda-1})$ некоторый однозначно-определенный элемент $[\varphi]$.

Лемма 3. (см. [54]). Пусть $M' = M_{\lambda-1} \cup_{\varphi} D^{\lambda} \times D^{n-\lambda} \cup_{\psi} D^{\lambda+1} \times D^{n-\lambda-1}$, где $\lambda \geq 2$, $n - \lambda \geq 2$. Существует такое вложение $\tilde{\psi} : \partial D^{\lambda} \times D^{n-\lambda} \rightarrow \partial M_{\lambda-1}$, изотопное вложению ψ , что $[\tilde{\psi}] = [\varphi] + [\psi]$. При этом мы считаем, что образы отображений φ и $\tilde{\psi}$ не пересекаются. Аналогично существует такое отображение $\tilde{\psi}$, что $[\tilde{\psi}] = [\psi] - [\varphi]$. Следовательно, индексы пересечения новой ручки, построенной по вложению $\tilde{\psi}$, получаются сложением (вычитанием) исходных индексов. При этой операции сложения ручек матрица гомоморфизма $\partial_{\lambda} : H_{\lambda}(M_{\lambda}, M_{\lambda-1}) \rightarrow H_{\lambda-1}(M_{\lambda-1}, M_{\lambda-2})$ подвергается элементарному преобразованию.

Переходим к непосредственному доказательству теоремы 1. Пусть f — некоторая правильная функция Морса на M^n и $\xi[f]$ — некоторое градиентно-подобное векторное поле. Построим соответствующее разложение многообразия в сумму ручек. Из условий теоремы 1 вытекает, что в этом разложении отсутствуют ручки индексов 1, $n-1$ и имеется по одной ручке индекса 0 (минимум функции) и индекса n (максимум функции). Рассмотрим цепочку включений $D^n = M_0 \subset M_2 \subset M_3 \subset \dots \subset M_{n-2} \subset M_n = M^n$. Строим соответствующий цепной комплекс $\{C, \partial\}$ и на основании леммы 1 получаем, что $H_i(C, \partial) = H_i(M^n, \mathbb{Z})$. Рассмотрим ручки индексов 2 и 3 и гомоморфизм $\partial_3 : H_3(M_3, M_2, \mathbb{Z}) \rightarrow H_2(M_2, M_0, \mathbb{Z})$, который задается матрицей индексов пересечения соответствующих правосторонних и левосторонних сфер. Используя описанный в лемме 3 процесс сложения ручек, приведем матрицу гомоморфизма ∂_3 к диагональному виду

$$\begin{pmatrix} \varepsilon_1 & \dots & \varepsilon_k & 0 \\ & & n_1 & \dots \\ & 0 & & n_l \end{pmatrix},$$

где $\varepsilon_i = \pm 1$, n_i делит n_{i+1} , $n_i \in \mathbb{Z}$.

Ясно, что ручки с индексом пересечения сфер, равным ε_i , вклад в группу гомологий не дают, и на основании леммы 2 эти пары ручек можно уничтожить. Ручки индекса 2, которые не пересекаются с ручками индекса 3 и соответствуют свободным образующим группы $H_2(M^n, \mathbb{Z})$, останутся. Останутся также ручки индексов 2 и 3, которые имеют пересечения друг с другом с индексами, равными n_i , что соответствует элементам конечного порядка в группе $H_2(M^n, \mathbb{Z})$. Используя двойственное разбиение на ручки (т. е. используя функцию $1-f$), анало-

тичными рассуждениями приведем к нужному виду ручки индексов $n-2$ и $n-3$. Рассмотрим затем ядро гомоморфизма ∂_3 . Это свободная абелева группа, порожденная ручками индекса 3, которые не пересекаются с ручками индекса 2. Уничтожим лишние ручки индекса 3 с помощью ручек индекса 4. Получаем, что $\partial_4 H_4(M_4, M_3, \mathbb{Z}) \subset \text{Ker } \partial_3$. Приведем затем матрицу гомоморфизма ∂_4 к диагональному виду и аналогичными рассуждениями докажем, что имеется $p^3 + q^3$ не уничтожаемых ручек. Таким образом, после двух шагов мы получили разложение многообразия M^n , имеющее $p^2 + q^2$ ручек индекса 2 и $p^3 + q^3 + q^2$ ручек индекса 3. Итерировав этот процесс, можно последовательно уничтожить лишние ручки остальных индексов. Теорема доказана.

Теорема 2 (см. [53]). Пусть M^n — односвязное гладкое замкнутое многообразие размерности $n \geq 6$. Пусть f и g — две точные функции Морса на многообразии M . Тогда существует эффективно вычислимое препятствие к существованию диффеоморфизма $h: M \rightarrow M$, такого, что $f = gh$.

Условие односвязности многообразия M^n является весьма существенным для справедливости теорем 1 и 2. Простые примеры показывают, что в неодносвязном случае минимальное число критических точек функции Морса на M^n зависит по крайней мере от гомологий универсального накрывающего многообразия $\tilde{M}^n$ и от некоторых тонких топологических инвариантов. Кроме того, оказывается (см. [53]), что появляется некоторая неоднозначность в построении точных функций Морса (в тех случаях, когда они существуют); см. также [52]. В некоторых случаях тем не менее удастся доказать существование точной функции Морса и вычислить число ее критических точек в каждой размерности. Мы приведем здесь только некоторые из этих результатов. Полный обзор см. в [53].

Теорема 3 (см. [53]). Пусть M^{n+k} — гладкое многообразие размерности $n \geq 6$ и пусть $k \geq 0$, $\pi_1(M^{n+k}) = \mathbb{Z} \oplus \dots \oplus \mathbb{Z}$ (k раз). Тогда на многообразии M^{n+k} существует точная функция Морса. Кроме того, число N_λ критических точек индекса λ у этой функции равно $p^{\lambda-k}(1+p)^k + q^{\lambda-k-1}(1+q)^{k+1}$, где p^i, q^i после раскрытия скобок (в соответствии с обычными алгебраическими правилами) понимаются как минимальное число свободных образующих в группе гомологий $H_i(\tilde{M}^n, \mathbb{Z})$ и как минимальное число образующих в подгруппе конечного порядка в группе $H_i(\tilde{M}^n, \mathbb{Z})$ соответственно. Здесь $\tilde{M}^n$ — универсальное накрывающее многообразие для многообразия M^n .

При $k=0$ получается теорема 1 (Смейл). При $k=1$ получаем, $N_\lambda = p^\lambda + p^{\lambda-1} + q^\lambda + 2q^{\lambda-1} + q^{\lambda-2}$. При $k=2$ имеем: $N_\lambda = p^\lambda + 2p^{\lambda-1} + p^{\lambda-2} + q^\lambda + 3q^{\lambda-1} + 3q^{\lambda-2} + q^{\lambda-3}$ и т. д. Таким образом, по мере роста фундаментальной группы многообразия в формировании минимального числа критических точек данного индекса λ принимают участие группы гомологий многообразия из все более широкого диапазона размерностей.

§ 2. Многозначный аналог теории Морса

Настоящий параграф посвящен изложению некоторых аспектов аналога теории Морса для многозначных функций и функционалов. Сначала мы кратко изложим новую теорию, построенную С. П. Новиковым в [59, 60, 61].

Пусть M — некоторое многообразие, конечномерное или бесконечномерное, на котором задана внешняя замкнутая дифференциальная 1-форма ω , т. е. $d\omega=0$. Интегралы этой формы по путям в многообразии M определяют некоторый «многозначный функционал» S на пространстве всех путей. Такие функционалы, естественно, возникают в физических задачах, поэтому их изучение представляет самостоятельный интерес [60]. В частности, сразу возникает следующий вопрос: можно ли построить аналог теории Морса для многозначных функционалов (функций), т. е. найти связи чисел стационарных (критических) точек (где $dS=0$) различных индексов с топологией многообразия M ?

Естественные примеры многозначных функционалов были рассмотрены в [61] при исследовании периодических решений уравнений типа Кирхгофа, волчка в гравитационном поле и др. Наиболее общий класс примеров многозначных функционалов может быть определен так. Рассмотрим два гладких многообразия N^q и M^n , где многообразие N^q предполагается компактным, и пусть S_0 — какой-либо функционал на пространстве F гладких отображений $f: N \rightarrow M$. Зададим на многообразии M^n замкнутую $(q+1)$ -мерную дифференциальную форму ω и континуальное покрытие $M^n = \bigcup_i U_i$ открытыми множествами U_i такое, что выполняются следующие свойства:

1) Для любого отображения $f: N \rightarrow M$ образ многообразия N , т. е. $f(N)$, целиком попадает в какое-то множество U_i . Номер i зависит, вообще говоря, от отображения f .

2) Форма ω точна на множестве U_i , т. е. $\omega = d\psi_i$.

Тогда, естественно, возникает некоторый многозначный функционал S на бесконечномерном пространстве F отображений $f: N \rightarrow M$; а именно для любого отображения $f: N \rightarrow U_i$ положим

$$S_i(f) = S_0(f) + \int_{f(N)} \psi_i. \quad (*)$$

Тогда функционал S_i можно рассматривать как «ветвь» многозначного функционала S , определенную на некотором открытом подмножестве в пространстве отображений. Функционалы S_i являются в этом смысле «локальными функционалами». Склеивая их, мы можем построить некоторый глобально определенный, но, вообще говоря, многозначный функционал S .

Если мы рассмотрим пересечение множеств $U_i \cap U_j$ и если образ $f(N)$ целиком содержится в этом пересечении, то $S_i(f) = S_j(f) = \int_{f(N)} (\psi_i - \psi_j)$, где $d(\psi_i - \psi_j) = 0$, т. е. $\psi_i - \psi_j$ является замкнутой q -формой на пересечении $U_i \cap U_j$. Это последнее ус-

ловие означает в действительности, что 1-форма δS корректно определена на всем бесконечномерном функциональном многообразии F . В самом деле, $\delta(S_i - S_j)(f) = \int d(\psi_i - \psi_j) = 0$, т. е. 1-форма δS одинаково вычисляется как в «карте» U_i , так и в «карте» U_j . Ясно, далее, что многозначность функционала S на пространстве F имеет своим первоисточником возможную негомотопность разных путей, соединяющих две фиксированные точки в многообразии F . Поэтому, переходя от многообразия F к некоторому накрывающему многообразию $\tilde{F}$ (фундаментальная группа которого «меньше» фундаментальной группы F), мы получаем на накрытии $\tilde{F}$ 1-форму, являющуюся дифференциалом некоторой однозначной функции на накрывающей $\tilde{F}$. Подводя итоги, сформулируем следующую лемму:

Лемма 1 (см. [60]). Набор локальных функционалов $S_i(f)$ определяет замкнутую 1-форму δS на функциональном пространстве F допустимых отображений $f: N \rightarrow M$. Эта 1-форма δS определяет многозначный функционал S на пространстве отображений F . Этот функционал является однозначным на каком-то бесконечнолистном накрытии $\tilde{F} \rightarrow F$.

Рассмотрим следующий пример. Пусть $N^1 = S^1$ является окружностью, тогда бесконечномерное многообразие F является пространством всех замкнутых путей (с параметризацией), в многообразии M . Если многообразие M односвязно, то пространство F связно. В качестве исходного функционала S_0 можно рассмотреть обычный функционал длины, определенный на путях в M , т. е. на пространстве отображений $S^1 \rightarrow M$. Этот функционал порождает некоторый (вообще говоря, многозначный) функционал S (см. выше).

Теорема 1 (см. [59]). Пусть M^n — односвязное полное риманово многообразие (т. е. снабженное положительно-определенной метрикой) и S_0 — функционал длины, определенный на пространстве F замкнутых путей в многообразии M . Тогда для любой замкнутой 1-формы ω на многообразии M^n функционал S вида (), однозначный или многозначный, имеет хотя бы одну периодическую экстремаль (критическую точку функционала на многообразии F).*

Перейдем к вопросу о связи гомологий многообразия с числом критических точек многозначных функций. Пусть M^n — конечномерное многообразие и ω — замкнутая 1-форма на M . Эта форма определяет элемент Ω группы когомологий $H^1(M, \mathbb{R})$. Поскольку $H^1(M, \mathbb{R}) \approx H_1(M, \mathbb{R})$, то элемент Ω определяет некоторый одномерный цикл с вещественными коэффициентами разложения по целочисленному базису в группе $H_1(M, \mathbb{Z}) \otimes \mathbb{R} = H_1(M, \mathbb{R})$. Следовательно, существует некоторое бесконечнолистное накрытие $p: \tilde{M} \rightarrow M$ такое, что форма $p^*\omega$ является точной на $\tilde{M}$, т. е. дифференциалом некоторой функции S на этом накрытии: $p^*\omega = dS$. Простейший пример: стандартная 1-форма $\omega = d\varphi$ на плоскости с выколотым началом координат, тогда накрытие p является римановой поверх-

ностью логарифма. Функцию S , корректно определенную на накрытии, естественно назвать «многозначной функцией» на исходном многообразии M^n . Будем предполагать в дальнейшем, что все ее критические точки невырождены либо образуют невырожденные критические подмногообразия в объемлющем многообразии. Будем предполагать также, что для функционала S корректно определен градиентный спуск, т. е. что на многообразии $\tilde{M}$ любое компактное множество при спуске вниз вдоль интегральных траекторий поля $\text{grad } S$ либо повисает на какой-либо критической точке (т. е. деформирующееся множество «цепляет» какой-то своей точкой критическую точку), либо последовательно проходит все поверхности уровня S , двигаясь «вниз». Обозначим через $N_\lambda(S)$ (или через $N_\lambda(\omega)$) число критических точек с индексом Морса λ у функции S , где $p^*\omega = dS$. Как и в случае однозначных функций, возникает вопрос: как оценить число критических точек многозначной функции S (т. е. замкнутой 1-формы ω) индекса λ через топологические инварианты многообразия?

В группе $H_1(M, \mathbf{Z})$ можно выбрать такой базис $(\gamma_1, \dots, \gamma_k, \gamma_{k+1}, \dots, \gamma_q)$, что для интегралов формы ω по одномерным циклам γ_i выполняются следующие равенства:

$$\int_{\gamma_j} \omega = \begin{cases} 0, & \text{если } j \geq k+1, \\ \kappa_j \neq 0, & \text{если } j \leq k, \end{cases}$$

причем все числа κ_j при $j=1, \dots, k$ линейно независимы с рациональными (или целыми) коэффициентами. Число $k-1$ называется степенью иррациональности формы ω . Группа монодромии (группа сдвигов) минимального накрытия $p: \tilde{M} \rightarrow M$, на котором прообраз формы ω превращается в дифференциал однозначной функции $p^*\omega = dS$, изоморфна группе $\mathbf{Z}^k$. Эта группа является свободной абелевой, и ее образующие $t_1, \dots, t_k$ реализуются как сдвиги на накрытии $\tilde{M}$, т. е. $t_j: \tilde{M} \rightarrow \tilde{M}$. Наиболее простым и в то же время интересным случаем является $k=1$, когда форма ω (возможно, после умножения на некоторое вещественное число) совпадает с некоторым элементом целочисленной группы когомологий $[\omega] \in H^1(M, \mathbf{Z})$. В нашем случае величина $\exp(2\pi i S)$ является однозначной комплекснозначной функцией, следовательно, корректно определено однозначное отображение $f = \exp(2\pi i S): M^n \rightarrow S^1$. В [60] изучены свойства критических точек таких отображений и их связь с топологией многообразия.

Ясно, что если построенное выше отображение f не имеет критических точек, то f определяет некоторое локально-тривиальное расслоение $f: M^n \rightarrow S^1$ с базой $B=S^1$. Поскольку $k=1$, то накрытие $p: \tilde{M} \rightarrow M$ является $\mathbf{Z}$ -циклическим, т. е. его слоем является группа целых чисел $\mathbf{Z}$. Опишем более подробно процесс построения этого накрытия. Пусть D — оператор двойственности Пуанкаре, устанавливающий изоморфизм между группами $H^1(M^n, \mathbf{Z})$ и $H_{n-1}(M^n, \mathbf{Z})$. Применяя его к одномерно-

му коциклу $[\omega]$, мы получаем элемент $D[\omega] \in H_{n-1}(M^n, \mathbb{Z})$. Реализуем цикл $D[\omega]$ подмногообразием K^{n-1} в M^n . Разрезая многообразие M^n по этому циклу-подмногообразию K^{n-1} , мы получаем пленку W^n с двумя краями $\partial W = K_0^{n-1} \cup K_1^{n-1}$, диффеоморфными K^{n-1} . Возьмем бесконечное число экземпляров этой пленки $W_i \approx W$ с границами $\partial W_i = K_{i,0} \cup K_{i,1}$, где $K_{i,\alpha}$ диффеоморфны K^{n-1} . Склеим их друг с другом по краям согласно

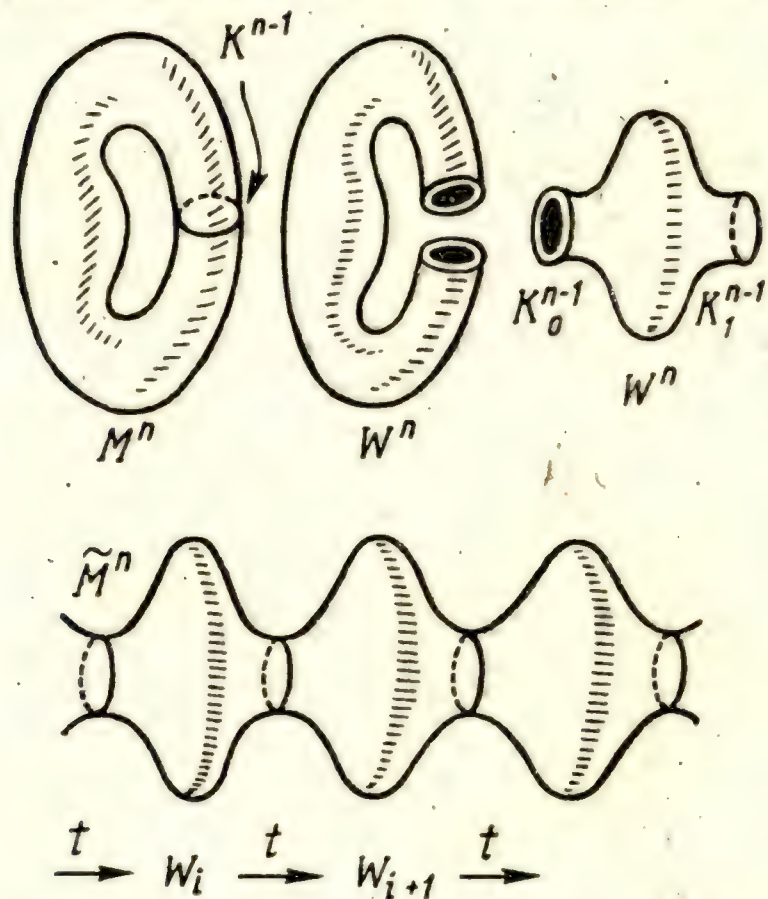


Рис. 3

правилу, указанному следующей нумерацией: $\tilde{M} = \bigcup_i W_i$, $K_{i+1,0} \approx K_{i,1}$, $-\infty < i < +\infty$ (рис. 3). В результате получается «бесконечная в обе стороны труба», составленная из бесконечного числа склеенных друг с другом экземпляров пленки W . Ясно также, что функция S корректно определена на этой трубе $\tilde{M}$, причем $\tilde{M}$ является накрытием над исходным многообразием M . Взятие этого накрытия разрушило коцикл $[\omega]$. Рассмотрим теперь поверхности уровня функции S . В силу выбора подмногообразия K , реализующего цикл $D[\omega]$, можно считать, что $K = K_{0,0}$ является одной из поверхностей уров-

ня функции S , т. е. полным прообразом точки при отображении $f = \exp(2\pi i S)$ многообразия $\tilde{M}$ в окружность. Из определения накрытия следует, что оператор t монодромии действует следующим образом на накрытии: $t: W_i \rightarrow W_{i+1}$, $K_{i,0} \rightarrow K_{i,1} = K_{i+1,0}$ (рис. 3).

Согласно классической теории Морса, гладкая функция Морса, заданная на конечномерном многообразии $\tilde{M}$, порождает при некоторых естественных условиях клеточный комплекс, являющийся разбиением многообразия $\tilde{M}$ в сумму клеток. Однако в рассматриваемом случае не выполняется важное требование, выполнение которого необходимо для справедливости обычной теории Морса. А именно для корректной реализации перестроек Морса необходимо требовать, чтобы каждая замкнутая область меньших значений $\{S \leq a\}$ (для всех a) являлась бы относительно компактной. Только при выполнении этого условия можно гарантировать, что область значений $\{S \leq a + \varepsilon\}$, где a — некоторое критическое значение для функции S , получается из области $\{S \leq a - \varepsilon\}$ приклейкой некоторого числа ручек. В рассматриваемом случае области меньших значений относительно компактными не являются. Однако и здесь из каждой критической точки индекса λ для функции S выходит «вниз» по траекториям градиентного спуска поверхность «наискорейшего спуска» — аналог сепаратрисного диска размерности λ в клас-

сическом случае. Эту поверхность (или ее достаточно малое возмущение, если это необходимо) можно считать «клеткой» размерности λ .

Однако, в отличие от классического случая, эта «клетка» может тянуться по уровням функции S до $-\infty$, и, следовательно, в ее алгебраическую границу может входить бесконечное число таких же «клеток» размерности $\lambda-1$, соответствующих критическим точкам индекса $\lambda-1$.

При сдвиге $t: \tilde{M} \rightarrow \tilde{M}$ функция S переходит в себя с добавлением некоторой постоянной. При этом критические точки функции переходят при сдвиге снова в критические точки. Следовательно, мы получаем:

1) Каждая критическая точка функции S определяет свободную образующую (реализованную «бесконечной клеткой», размерность которой равна индексу критической точки) в комплексе цепей; 2) граница каждой такой клетки является линейной комбинацией (возможно, бесконечной) клеток этого клеточного комплекса, лежащих «ниже» по уровням функции S , т. е. уходящих в бесконечность только в одну сторону в накрытии-«трубе» $\tilde{M}$; 3) все клетки построенного комплекса получаются из конечного числа базисных клеток всевозможными сдвигами на элементы t^m группы $\mathbf{Z}$, действующей на накрытии $\tilde{M}$ описанным выше образом.

Опишем некоторую алгебраическую конструкцию, на основе которой можно, как выясняется (см. [59, 60]), можно получить неравенства типа Морса для многозначных функций. Рассмотрим кольцо L , состоящее из рядов Лорана следующего вида:

$\sum_{j > \text{const} > -\infty} m_j t^j$. Здесь m_j — целые числа, обращающиеся

в ноль для всех достаточно больших отрицательных номеров j . Клеточный комплекс, порожденный многозначной функцией на многообразии M или однозначной функцией S на его накрытии $\tilde{M}$, можно рассматривать как свободный комплекс C , состоящий из L -модулей C_i с конечным числом образующих. Здесь мы используем то обстоятельство, что число критических точек на каждом многообразии W_i конечно. Комплекс C имеет следующий вид:

$$0 \rightarrow C_n \xrightarrow{\partial} C_{n-1} \xrightarrow{\partial} \dots \xrightarrow{\partial} C_1 \xrightarrow{\partial} C_0 \rightarrow 0,$$

где граничный оператор ∂ является гомоморфизмом L -модулей. В отличие от обычной теории Морса здесь возможна ситуация, когда $C_0 = C_n = 0$. В том случае, когда отображение $f: M^n \rightarrow S^1$ определяет локально тривиальное расслоение с базой окружность, на M существует 1-форма ω , вообще не имеющая критических точек. В качестве такой формы можно взять прообраз формы $d\varphi$ на S^1 . В этом случае комплекс тривиален, т. е. $C_i = 0$ для всех i . Тем не менее, как и в классическом случае,

комплекс C позволяет вычислять гомологические характеристики многообразия.

Лемма 2 (см. [60]). *Гомологии комплекса L -модулей C , порожденного любой гладкой замкнутой 1-формой ω с невырожденными критическими точками и корректно определенным градиентным спуском, гомотопически инвариантны.*

Следовательно, инварианты групп гомологий описанного комплекса могут быть использованы для получения аналогов неравенств Морса в случае многозначных функций, порождающих (см. выше) отображения в окружность $f = \exp(2\pi i S) : M^n \rightarrow S^1$. Легко видеть, что подмодули свободных модулей C_i свободны. Следовательно, в модулях циклов $Z_i = \text{Ker } d \subset C_i$ и границ $B_i = \text{Im } d \subset C_i$ можно выбрать свободные базисы. Разность рангов A и B этих свободных подмодулей назовем числом Бетти и обозначим его через $p^i(M^n, a)$, где $a = [\omega] \in H^1(M^n, \mathbb{Z})$. Ясно, что в классическом случае эти числа превращаются в обычные числа Бетти $p^i(M^n)$ (см. § 1 настоящей главы). Аналоги $q^i(M^n, a)$ классических чисел кручения $q^i(M^n)$ определяются так. Выберем в модуле циклов Z_i свободный базис $e_1, \dots, e_A$, а в подмодуле границ B_i выберем свободный базис $e'_1, \dots, e'_B$, где $A - B = p^i$, причем всегда можно выбрать эти базисы так, что будут выполняться следующие свойства:

$$1) \quad e'_j = (n_j + \sum_{k \geq 1} n_{jk} t^k) e_j + \sum_{i > B} q_{ij}(t) e_i;$$

2) число n_j делится на число n_{j+1} , 3) степени всех членов ряда $q_{ij}(t)$ неотрицательны, 4) числа $q_{ij}(0) \neq 0$ и делятся на n_j для всех i, j (если ряд не обращается в ноль тождественно).

Общее число индексов j таких, что $n_j \neq 1$, называется числом кручения и обозначается через $q^i(M^n, a)$. Число $p^i + q^i$ совпадает с минимальным числом образующих модуля $H_i = Z_i/B_i$, что является естественным аналогом классического случая.

Теорема 2 (см. [60]). *Справедливы аналоги неравенств Морса для чисел $N_\lambda(S)$ (или $N_\lambda(\omega)$) критических точек индекса λ для отображения $f = \exp(2\pi i S)$ многообразия M^n в окружность S^1 , или замкнутой 1-формы ω , где $a = [\omega] \in H^1(M^n, \mathbb{Z})$, т. е. форма ω (возможно после умножения на некоторое число) совпадает с некоторым элементом целочисленной группы одномерных когомологий. Эти аналоги неравенств имеют вид:*

$$N_\lambda(S) \geq p^\lambda(M^n, a) + q^\lambda(M^n, a) + q^{\lambda-1}(M^n, a).$$

По форме эти неравенства совпадают с классическими (см. предыдущий параграф), однако входящие в них топологические инварианты имеют более глубокий и сложный геометрический смысл.

Если многообразие M^n имеет фундаментальную группу, изоморфную группе $\mathbb{Z}$, и если 1-форма ω реализует коцикл, соответствующий образующей этой группы, то накрытие $\tilde{M}^n$ одно-

связно, и, следовательно, возникает вопрос о точности аналогов неравенств Морса, полученных в теореме 2, в том же смысле, в каком понимается их точность в классическом случае (см. предыдущий параграф, теорему Смейла). Другими словами, возникает вопрос о том, можно ли построить такую многозначную функцию, чтобы неравенство из теоремы 2 превратилось в точное равенство, т. е. всегда ли существует точная многозначная функция Морса. Построение «минимальной» (в этом смысле) 1-формы ω требует выбора в некотором смысле «минимального» начального подмногообразия $K^{n-1} \subset M^n$, реализующего цикл, двойственный коциклу ω . Оказывается, как доказали М. Ш. Фарбер и В. В. Шарко, для указанного выше случая точная многозначная функция Морса всегда существует.

Аналогичная теория многозначных функций и неравенств Морса строится (см. [60]) и для случая, когда степень иррациональности формы ω больше нуля, т. е. когда $k > 1$ (см. выше). Общая схема конструкции остается прежней, но при этом появляются некоторые новые интересные эффекты, связанные с кольцами рядов Лорана от многих переменных.

Теперь мы опишем постановку другой задачи, сформулированной автором настоящей книги и тесно связанной с многозначными отображениями, характеризующими некоторые функционалы. В четвертой главе мы рассмотрим двумерные минимальные поверхности в евклидовом пространстве $\mathbb{R}^3$, затачивающие замкнутые контуры, гомеоморфные окружности. При деформации контура поверхность изменяется; в частности, изменяется ее площадь. Итак, можно рассмотреть график площади минимальной поверхности, определенный на пространстве всех граничных контуров в $\mathbb{R}^3$. Так как на одном и том же контуре могут повисать, вообще говоря, несколько минимальных поверхностей, то график площади оказывается многозначной функцией контура. Теперь сформулируем общую задачу. Рассмотрим пространство $C^\infty(M, N)$ гладких отображений $f: M \rightarrow N$, где N — фиксированное риманово многообразие, а M — всевозможные гладкие римановы многообразия с непустым краем ∂M . На пространстве $C^\infty(M, N)$ рассмотрим какой-либо функционал $F(f)$, задающийся так: $F(f) = \int_M L(f) d\sigma$, где L — лагранжиан, зависящий от отображения f и от его производных, а $d\sigma$ — форма риманова объема на многообразии M . В качестве примера можно рассмотреть функционал vol объема (или площади) образа $f(M)$ в N . Рассмотрим уравнения Эйлера—Лагранжа $I(f) = 0$ для функционала F . Экстремали f_0 функционала F — это решения уравнения Эйлера—Лагранжа $I(f_0) = 0$. Например, экстремалими функционала объема являются локально минимальные поверхности. Двумерные минимальные поверхности моделируются мыльными пленками в $\mathbb{R}^3$, затачивающими замкнутые проволоочные контуры. Для задания экстремали f_0 нужно задать граничные (краевые) условия, т. е. задать

отображение $\gamma: \partial M \rightarrow N$. Например, для задания мыльной пленки нужно фиксировать в $\mathbb{R}^3$ какой-либо замкнутый контур γ , т. е. фиксировать отображение окружности S^1 в $\mathbb{R}^3$. Множество всех граничных условий естественно отождествляется с пространством $C^\infty(\partial M, N)$ отображений $\gamma: \partial M \rightarrow N$. Задание того или иного граничного условия определяет, вообще говоря, несколько экстремалей f_0 функционала F , т. е. несколько решений уравнений Эйлера — Лагранжа таких, что $f_0|_{\partial M} = \gamma$. Обозначим через K_γ множество всех экстремалей с данным граничным условием γ . Вычисляя для каждой из них значение функционала F , мы получаем множество чисел $\{F(f_0)\}$. В результате мы получаем некоторое многозначное отображение (функцию) $\gamma \rightarrow \{F(f_0)\}$, где $f_0 \in K_\gamma$. Эта функция определена на пространстве $C^\infty(\partial M, N)$ и имеет, вообще говоря, точки ветвления, другие особенности. Разные функционалы (и их уравнения Эйлера — Лагранжа) будут характеризоваться, вообще говоря, разными типами особенностей. Общая задача: какие типы особенностей характеризуют те или иные уравнения Эйлера — Лагранжа? При конкретном исследовании многозначных функций вида $\gamma \rightarrow \{F(f_0)\}$ полезно на первых этапах упрощать задачу, рассматривая не все пространство $C^\infty(\partial M, N)$, а лишь конечномерные подпространства в $C^\infty(\partial M, N)$, на которые нужно ограничивать функцию $\gamma \rightarrow \{F(f_0)\}$. В гл. 4 мы, например, изучим график этой функции, ограниченной на двумерное подмногообразие в $C^\infty(\partial M, N)$ для функционала площади двумерных мыльных пленок. Более простым вариантом общей задачи является случай функционала длины (или действия), определенного на пространстве кусочно-гладких путей, соединяющих пару точек на гладком связном римановом многообразии. В результате мы получаем многозначную функцию на $N \times N$, где N — заданное многообразие. Два листа этой многозначной функции пересекаются, если появляются две геодезические одинаковой длины, соединяющие пару точек (но геометрически различные геодезические). Здесь экстремалами функционала являются геодезические с фиксированными концами. Уже для простого случая плоского тора эта многозначная функция имеет любопытную структуру.



§ 1. Гомеоморфизмы двумерных поверхностей

1. Группа гомеотопий и классификация трехмерных многообразий рода 1.

В настоящей главе мы будем в некоторой степени опираться на материал, изложенный, например, в [34]. Обозначим через Π трехмерное многообразие, с краем, ограниченное в $\mathbb{R}^3$ стандартно вложенной двумерной сферой M_g^2 с g ручками (рис. 1). На рисунке показан случай $g=3$. Будем для сокращения называть многообразие Π полным (заполненным) кренделем рода g . Рассмотрим два экземпляра полного кренделя

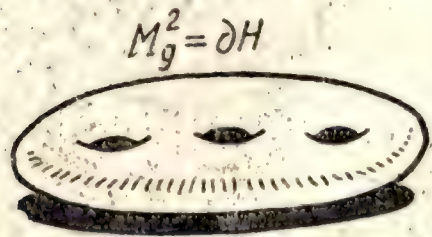


Рис. 1

Π_1 и Π_2 . Хорошо известно, что если рассмотреть произвольный гомеоморфизм $\alpha: \partial\Pi_1 \rightarrow \partial\Pi_2$, то можно построить некоторое трехмерное замкнутое компактное связное ориентируемое многообразие $M(\alpha)$. Для этого достаточно склеить два полных кренделя Π_1 и Π_2 по гомеоморфизму их границ. Верно и обратное: любое трехмерное многообразие указанного вида может быть представлено в виде

$M(\alpha)$ для некоторого гомеоморфизма $\alpha: \partial\Pi_1 \rightarrow \partial\Pi_2$ поверхностей некоторого рода g . Такое представление неоднозначно.

Среди множества всех гомеоморфизмов $\alpha: \partial\Pi_1 \rightarrow \partial\Pi_2$ выделяются гомеоморфизмы, которые продолжаются на внутренность, т. е. продолжаются до гомеоморфизма полных кренделей $\hat{\alpha}: \Pi_1 \rightarrow \Pi_2$.

Лемма 1. Если $\alpha_1, \alpha_2: \partial\Pi_1 \rightarrow \partial\Pi_2$ — гомеоморфизмы, продолжающиеся на внутренность сферы с g ручками, то трехмерные многообразия $M(\alpha)$ и $M(\alpha_2\alpha_1)$ гомеоморфны для любого гомеоморфизма $\alpha: \partial\Pi \rightarrow \partial\Pi$.

Доказательство. Определим непрерывное отображение $\varphi: M(\alpha_2\alpha_1) \rightarrow M(\alpha)$ следующим образом:

$$\varphi(x) = \begin{cases} \hat{\alpha}_1(x), & \text{если } x \in \Pi_1, \\ \hat{\alpha}_2(x), & \text{если } x \in \Pi_2. \end{cases}$$

Здесь через $\hat{\alpha}_i$ обозначены продолжения гомеоморфизмов α_i . Легко проверяется, что отображение φ — гомеоморфизм.

Напомним, что два гомеоморфизма $\alpha_1, \alpha_2: \partial\Pi \rightarrow \partial\Pi$ называются изотопными, если они гомотопны и если существует гомо-

топия f_t , связывающая α_1 и α_2 , т. е. $f_0 = \alpha_1$, $f_1 = \alpha_2$, и такая, что каждое отображение $f_t : \partial\Pi \rightarrow \partial\Pi$ является гомеоморфизмом.

Лемма 2. Если гомеоморфизм $\alpha : \partial\Pi \rightarrow \partial\Pi$ изотопен тождественному отображению $\partial\Pi$ на себя, то α продолжается на внутренность, т. е. до гомеоморфизма полного кренделя.

Доказательство вытекает из того, что тождественное отображение края полного кренделя на себя, очевидно, продолжается на внутренность.

Рассмотрим группу всех гомеоморфизмов двумерной поверхности $\partial\Pi$ рода g . Обозначим эту группу через H_g . В ней выделим подгруппу I_g тех гомеоморфизмов, которые изотопны тождественному.

Лемма 3. Группа I_g , состоящая из гомеоморфизмов двумерной поверхности $\partial\Pi$, изотопных тождественному, является нормальным делителем в группе всех гомеоморфизмов H_g .

Это утверждение очевидно. Следовательно, корректно определена фактор-группа $\Gamma_g = H_g / I_g$, которая называется группой гомеотопий двумерного кренделя $\partial\Pi$ рода g . Вычислим эту группу для случая $g=1$. Построим гомоморфизм ψ группы гомеотопий Γ_1 в группу всех автоморфизмов абелевой группы $\mathbf{Z} \oplus \mathbf{Z}$, реализованной в виде фундаментальной группы тора T^2 , т. е. $\mathbf{Z} \oplus \mathbf{Z} = \pi_1(S^1 \times S^1)$. Указанная группа автоморфизмов совпадает с группой обратимых целочисленных матриц порядка два с определителем, равным ± 1 , которую мы обозначим через $SL(2, \mathbf{Z})$. Гомоморфизм ψ определим так: каждый гомеоморфизм α поверхности индуцирует некоторый автоморфизм $\psi(\alpha)$ ее фундаментальной группы $\mathbf{Z} \oplus \mathbf{Z}$. Ясно, что гомеоморфизм, изотопный тождественному, индуцирует тождественный автоморфизм.

Теорема 1. Гомоморфизм $\psi : \Gamma_1 \rightarrow SL(2, \mathbf{Z})$ является изоморфизмом.

Доказательство. Образующими матрицами в группе $SL(2, \mathbf{Z})$ являются следующие:

$$\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Для доказательства эпиморфности отображения ψ достаточно представить эти матрицы как автоморфизмы целочисленной решетки $\mathbf{Z} \oplus \mathbf{Z}$, порожденные гомеоморфизмами тора. Ясно, что в качестве таких гомеоморфизмов нужно взять соответственно следующие отображения тора на себя: а) отражение тора относительно плоскости yz (рис. 2(а)), б) отражение тора относительно плоскости xz (рис. 2(б)), в) преобразование тора — скручивание на угол 2π , показанное на рис. 2(в) и называемое операцией Дена, г) преобразование, показанное на рис. 2(г) (также преобразование Дена). Поясним операции (в) и (г). Ввиду их однотипности достаточно прокомментировать одну из них, например, (в). Рассмотрим меридиан тора (вертикаль-

ное сечение), разрежем тор по этой кривой, затем закрутим правую часть получившегося цилиндра на угол 2π и снова приклеим его к левому краю цилиндра. Ясно, что при этом параллель, имевшая тип $(1,0)$, превратится в траекторию типа $(1,1)$, обходящую тор один раз вдоль параллели и один раз вдоль меридиана. Докажем мономорфность отображения ψ .

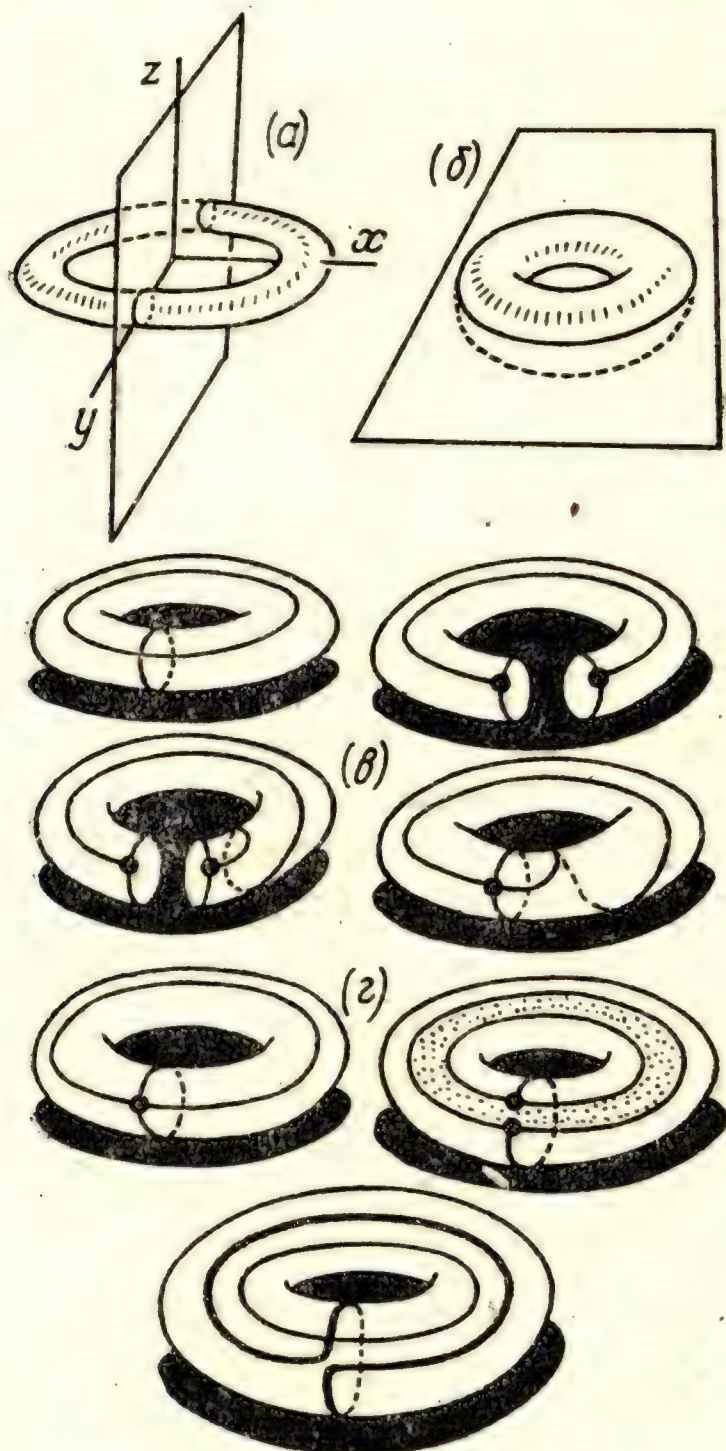


Рис. 2

Поскольку они гомотопны, то их индекс пересечения равен нулю. Определение и свойства индекса пересечения см., например, в [1, 4, 34]. Рассмотрим сначала тот случай, когда кривые пересекаются. В каждой точке определен знак пересечения. Суммируя эти ± 1 , мы и получаем индекс пересечения, равный нулю в силу гомотопности кривых. Пусть A и B — две соседние точки разных знаков, являющиеся точками пересечения γ_1 и γ_2 . Пусть l — дуга, соединяющая эти две соседние точки (рис. 3). Разрежем тор вдоль кривой γ_2 . Так как эта кривая не разбивает тор, то в результате разреза мы получим цилиндр, на одной из граничных окружностей которого расположены точки A и B . Так как дуга l ограничивает часть цилиндра, гомеоморфную диску, то можно с помощью изотопии «снять» дугу l и тем самым убрать две точки пересечения про-

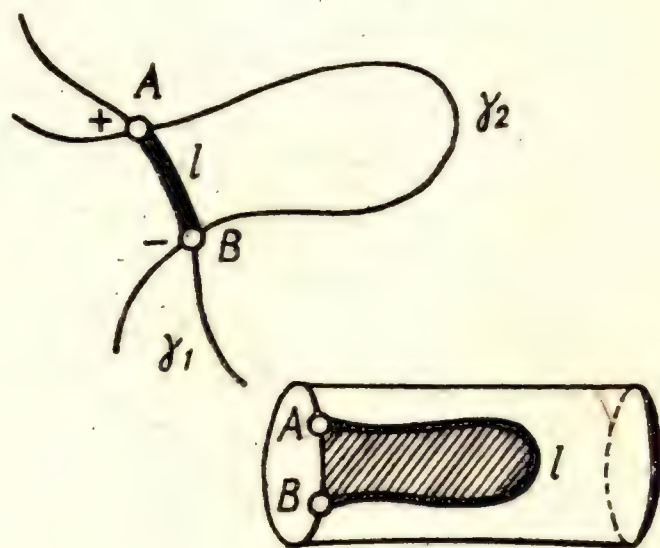


Рис. 3

Пусть дан некоторый гомеоморфизм тора такой, что индуцированный им автоморфизм фундаментальной группы является тождественным отображением. Нужно доказать, что гомеоморфизм изотопен тождественному.

Лемма 4. Если две замкнутые простые несамопересекающиеся кривые на торе не разбивают тор (в объединение двух непересекающихся областей) и гомотопны, то они изотопны.

Доказательство. Пусть γ_1 и γ_2 — две такие кривые.

тивоположных знаков. Итак, мы уменьшили число точек пересечения кривых. Продолжая этот процесс, мы в конце концов ликвидируем все точки пересечения обеих кривых. Таким образом, применяя соответствующую изотопию, мы можем считать, что кривые γ_1 и γ_2 не пересекаются. Так как они по-прежнему не разбивают тор, то, делая разрез по одной из них, например по γ_2 , мы получаем цилиндр, вокруг которого обвивается (без самопересечений) кривая γ_1 . Ясно, что существует изотопия, совмещающая γ_1 и γ_2 , что и требовалось.

Возвращаемся к доказательству теоремы. Рассмотрим две стандартные образующие группы $\pi_1(T^2)$ — параллель b и меридиан a . Фиксируем, например, меридиан a и рассмотрим его образ при гомеоморфизме α . По условию кривая $\alpha(a)$ принадлежит тому же гомотопическому классу, что и кривая a , а потому они гомотопны. Очевидно, обе кривые не разбивают тор. Следовательно, по лемме 4 существует изотопия, переводящая $\alpha(a)$ обратно в a . Аналогичным образом доказывается существование изотопии, возвращающей кривую $\alpha(b)$ в ее исходное положение b . При этом требуется некоторая аккуратность, чтобы опять не сдвинуть a . Таким образом, возвращая меридиан, а затем параллель в их первоначальное положение, получаем гомеоморфизм тора, тождественный на образующих a и b , т. е. на одномерном остове тора. Так как дополнение к объединению этих кривых на торе является диском, то завершение построения искомой изотопии на двумерном диске уже не составляет труда, так как на границе диска гомеоморфизм уже превращен (с помощью изотопии) в тождественный. Теорема доказана.

Таким образом, если Q_1 — множество всех трехмерных многообразий, представимых в виде $M(\alpha)$, где α — гомеоморфизм тора на себя (т. е. мы рассматриваем крендели рода 1), то определен эпиморфизм $B(2, \mathbf{Z}) \setminus \text{SL}(2, \mathbf{Z}) / B(2, \mathbf{Z}) \rightarrow Q_1$. Здесь через $B(2, \mathbf{Z})$ обозначена подгруппа матриц в группе $\text{SL}(2, \mathbf{Z})$, обладающих тем свойством, что соответствующие им гомеоморфизмы тора продолжаются на внутренность. Символ, включающий две косые черты, обозначает множество классов смежности, образованных правыми и левыми сдвигами на элементы подгруппы $B(2, \mathbf{Z})$. Здесь мы опираемся на лемму 1.

Лемма 5. Матрица $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbf{Z})$ принадлежит подгруппе $B(2, \mathbf{Z})$ в том и только в том случае, когда $c=0$.

Доказательство. Необходимость вытекает из того, что образ меридиана обязан стягиваться в полнотории, т. е. не должен наматываться на параллель. Достаточность следует из простого замечания, что на образ меридиана можно натянуть диск, и тогда на оставшуюся трехмерную клетку искомый гомеоморфизм продолжается по радиусам.

Рассмотрим подробнее операцию домножения целочислен-

ной матрицы справа и слева на матрицы вида $\begin{pmatrix} \pm 1 & \pm k \\ 0 & \pm 1 \end{pmatrix}$, т. е. на элементы подгруппы $B(2, \mathbf{Z})$. Напомним, что $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \pm 1$. Ясно, что указанная операция эквивалентна смене знаков строк и столбцов и добавлению первого столбца ко второму или второй строки к первой (с коэффициентом k).

Лемма 6. *Любая матрица из группы $SL(2, \mathbf{Z})$ приводится указанными левыми и правыми умножениями на матрицы из $B(2, \mathbf{Z})$ к одной из следующих матриц:* 1) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$; 2) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$;

3) $\begin{pmatrix} q & x \\ p & y \end{pmatrix}$, где $p \geq 2$, $0 < q < p$, $qy - px = 1$.

Доказательство. 1) Если исходная матрица такова, что в ее левом нижнем углу стоит ноль, то она имеет вид $\begin{pmatrix} \pm 1 & \pm k \\ 0 & \pm 1 \end{pmatrix}$ и, очевидно, приводится указанными операциями к виду $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. 2) Если в левом нижнем углу стоит 1, то сначала добиваемся (применением указанных операций) того, что на диагонали матрицы будут стоять нули. В этом случае автоматически в правом верхнем углу стоит ± 1 . 3) Матрицы, не удовлетворяющие условиям пунктов 1 и 2, очевидно, приводятся к виду 3.

Легко проверяется (предоставляем это читателю), что имеют место следующие гомеоморфизмы: если $\psi(\alpha) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, то $M(\alpha) = S^2 \times S^1$; если $\psi(\alpha) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, то $M(\alpha) = S^3$.

Если $\psi(\alpha) = \begin{pmatrix} q & x \\ p & y \end{pmatrix}$ есть матрица типа 3, то соответствующее многообразие $M(\alpha)$ обозначается $L(p, q)$ и называется линзой.

Теорема 2. *Линзы $L(p, q)$ и $L(p_1, q_1)$ гомеоморфны тогда и только тогда, когда выполняются соотношения: 1) $p = p_1$ и 2) $qq_1 \equiv \pm 1 \pmod{p}$ или $q \equiv \pm q_1 \pmod{p}$.*

В одну сторону эта теорема доказывается довольно просто. Докажем, что если выполнены условия 1 и 2, то линзы $L(p, q)$ и $L(p_1, q_1)$ гомеоморфны. Рассмотрим случай, когда $p = p_1$ и $q = \pm q_1 \pmod{p}$, т. е. $q = \pm q_1 + mp$ для некоторого целого m .

Тогда матрица $\begin{pmatrix} \pm q \pm mp & x_1 \\ p & y_1 \end{pmatrix}$, очевидно, получается из матрицы $\begin{pmatrix} q & x \\ p & y \end{pmatrix}$ описанными выше операциями. Так как эти опе-

рации не меняют тип многообразия, то соответствующие линзы гомеоморфны. Если $p=p_1$ и $qq_1=\pm 1+tp$, то, заменяя матрицу $\begin{pmatrix} q & x \\ p & y \end{pmatrix}$ на обратную, мы не меняем соответствующего многообразия, так как ясно, что $M(\alpha)=M(\alpha^{-1})$. В то же время обратная матрица имеет вид $\begin{pmatrix} y & -p \\ -x & q \end{pmatrix}$, т. е. элемент q заменился на ему «обратный» элемент y , так как $qy=1+px$. Доказательство обратного утверждения теоремы более нетривиально, и мы вынуждены его здесь опустить.

Тем самым лемма 6 и теорема 2 дают полное описание всех трехмерных многообразий, допускающих представление в виде $M(\alpha)$, где α — гомеоморфизм рода один, т. е. гомеоморфизм тора.

2. Система образующих группы гомеотопий полного кренделя.

Как мы уже видели, изучение группы Γ_g гомеотопий двумерного кренделя $\partial\Pi$ рода g позволяет получать информацию о трехмерных многообразиях, допускающих представление вида $M(\alpha)$, где α — некоторый гомеоморфизм поверхности $\partial\Pi$ рода g на себя. Более точно, пусть $K_g \subset H_g$ — подгруппа группы H_g гомеоморфизмов двумерной поверхности $\partial\Pi$, состоящая из продолжающихся на полный крендель Π гомеоморфизмов двумерного кренделя $\partial\Pi$ рода g . Пусть I_g' — группа гомеоморфизмов полного кренделя, изотопных тождественному. По лемме 2 получаем включение $\Gamma_g \subset K_g$, поэтому группа $G_g = K_g/I_g'$ содержится в группе $\Gamma_g = H_g/I_g$. Группу G_g можно интерпретировать как группу гомеотопий полного кренделя. Из леммы 1 получаем следствие: существует отображение

$$G_g \setminus \Gamma_g / G_g \rightarrow Q_g,$$

являющееся отображением «на», где через Q_g обозначено множество трехмерных многообразий вида $M(\alpha)$, где α — гомеоморфизм рода g , а $G_g \setminus \Gamma_g / G_g$ — множество двойных классов смежности, т. е. образованных гомеоморфизмами вида $\alpha_1 h \alpha_2$, ($\alpha_i \in G_g$, $h \in \Gamma_g$). С точки зрения задачи описания множества всех трехмерных многообразий значительный интерес представляет, следовательно, описание группы Γ_g и ее подгруппы G_g . Для удобства пользования введенными обозначениями приведем их таблицу:

полный крендель Π	его граница $\partial\Pi$	группа
K_g'	H_g	гомеоморфизмов
I_g'	I_g	гомеоморфизмов, изотопных тождественному
нет	K_g	гомеоморфизмов, продолжающихся на внутренность
$G_g = K_g/I_g'$	$\Gamma_g = H_g/I_g$	гомеотопий

Оказывается, что группа Γ_g порождена скручиваниями Дена вокруг кривых $s_i, l_i, 1 \leq i \leq g$ и $m_i, 1 \leq i \leq g-1$ (всего $3g-1$ кривых; рис. 4), а также симметриями относительно плоских сечений.

Обозначим через p какой-нибудь фиксированный гомеоморфизм кренделя $\partial\Pi$, неподвижный на кривых $s_i, 2 \leq i \leq g$ и переводящий кривую s_1 в кривую m_1 . Ясно, что такой гомеоморфизм существует и продолжается на внутренность. Рассмотрим

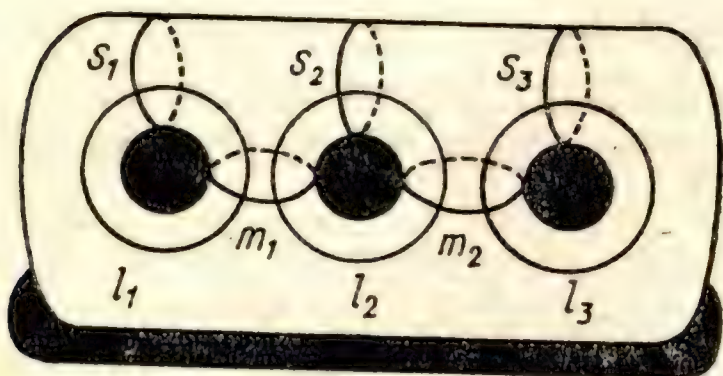


Рис. 4

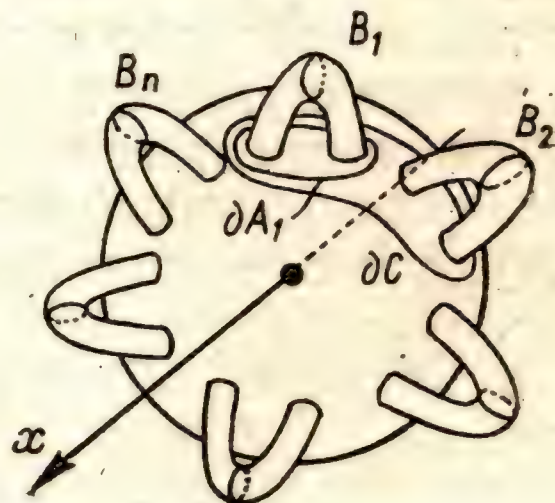


Рис. 5

группу S_g гомеоморфизмов полного кренделя, обладающих тем свойством, что они переводят множество кривых $s_i, 1 \leq i \leq g$ в себя, т. е. инвариантны на множестве кривых s_i . Ясно, что $S_g \subset K_g$. Образ этой группы при проекции $K_g \rightarrow G_g$ обозначим через $[S_g]$.

Теорема 3. *Группа G_g порождена всеми элементами группы $[S_g]$ и элементом $[p]$. Кроме того, группа G_g имеет конечное число образующих.*

Эта теорема доказана в 1973 г. С. В. Матвеевым. Доказательство можно также извлечь из публикации [35].

Оказывается, более детальный анализ позволяет упростить это описание группы. Представим полный крендель так, как это показано на рис. 5. Пусть a_1 — гомеоморфизм этого кренделя, являющийся поворотом кренделя на угол $2\pi/g$ вокруг оси x . Через a_2, a_3, a_4 обозначим гомеоморфизмы полного кренделя, возникающие в результате следующих операций: а) в случае a_2 и a_3 осуществим скручивание на угол π вокруг окружностей ∂A_1 и ∂C соответственно (рис. 5); б) в случае a_4 скрутим крендель на угол 2π вокруг окружности ∂B_1 . Затем положим $a_5 = p$, где гомеоморфизм p определен выше. Наконец, в качестве a_6 возьмем любой гомеоморфизм полного кренделя, меняющий ориентацию.

Теорема 4. *Группа гомеотопий полного кренделя порождена шестью элементами $a_1, a_2, \dots, a_6$. Группа сохраняющих ориентацию гомеотопий полного кренделя порождена пятью элементами $a_1, a_2, \dots, a_5$.*

§ 2. Алгоритм распознавания стандартной трехмерной сферы в классе диаграмм Хегора рода два

1. Графы Уайтхеда, кодирующие трехмерные многообразия. Операции индексов один и два.

Как уже было отмечено, любое трехмерное замкнутое многообразие можно представить (неоднозначно) в виде $M(\alpha)$, т. е. в виде склейки двух полных кренделей Π_1, Π_2 некоторого рода g по некоторому диффеоморфизму их границ $\alpha: \partial\Pi_1 \rightarrow \partial\Pi_2$. Представление многообразия в таком виде иногда называется диаграммой Хегора трехмерного многообразия. Фиксируем на двумерном кренделе набор стандарт-

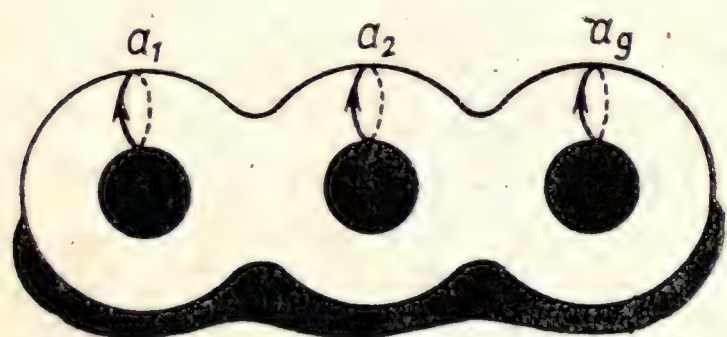


Рис. 6

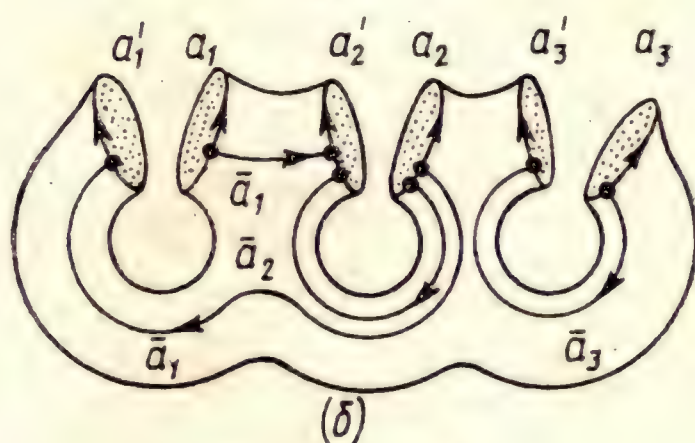
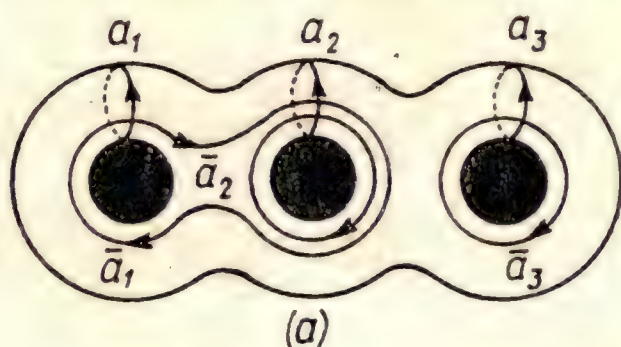


Рис. 7(a), (б)

ных меридианов $a_1, \dots, a_g$ (рис. 6). Рассмотрим их образы при диффеоморфизме α и обозначим их через $\bar{a}_1, \dots, \bar{a}_g$, $\bar{a}_i = \alpha(a_i)$. Без ограничения общности можно считать, что окружности a_i и $\bar{a}_j$ пересекаются трансверсально, в частности, в конечном числе точек.

Определение 1. Будем говорить, что система из $2g$ окружностей $a_1, \dots, a_g, \bar{a}_1, \dots, \bar{a}_g$ образует сеть s на двумерном кренделе. Число g назовем родом сети s . Назовем окружности a_i окружностями индекса один, а $\bar{a}_j$ — окружностями индекса два.

Эта терминология обязана своим происхождением имеющейся связи сетей с функциями Морса на трехмерных многообразиях. Анализ этой связи нам здесь не потребуется. Читатель, интересующийся этими вопросами, может обратиться к [40]. Из определения сети следует, что каждой сети соответствует замкнутое трехмерное многообразие, определенное с точностью до гомеоморфизма. Кроме того, каждое трехмерное многообразие M определяет бесконечное множество сетей, соответствующих всем диффеоморфизмам α таким, что $M = M(\alpha)$. В этом смысле множество сетей можно рассматривать как множество кодов, с помощью которых кодируются трехмерные многообразия. Отметим полезное для дальнейшего свойство сети. Разрежем крендель по всем окружностям индекса один. Ясно, что получится двумерная сфера с $2g$ дырками. Каждая из окружностей a_i порождает при разрезании кренделя два берега разреза, которые обозначим a_i и a_i' (с сохране-

нием ориентации). Аналогично, разрезая крендель по всем окружностям индекса два, мы также получаем сферу с $2g$ дырками. Будем считать, что на окружностях индекса один была фиксирована стандартная ориентация (рис. 6), тогда все ок-

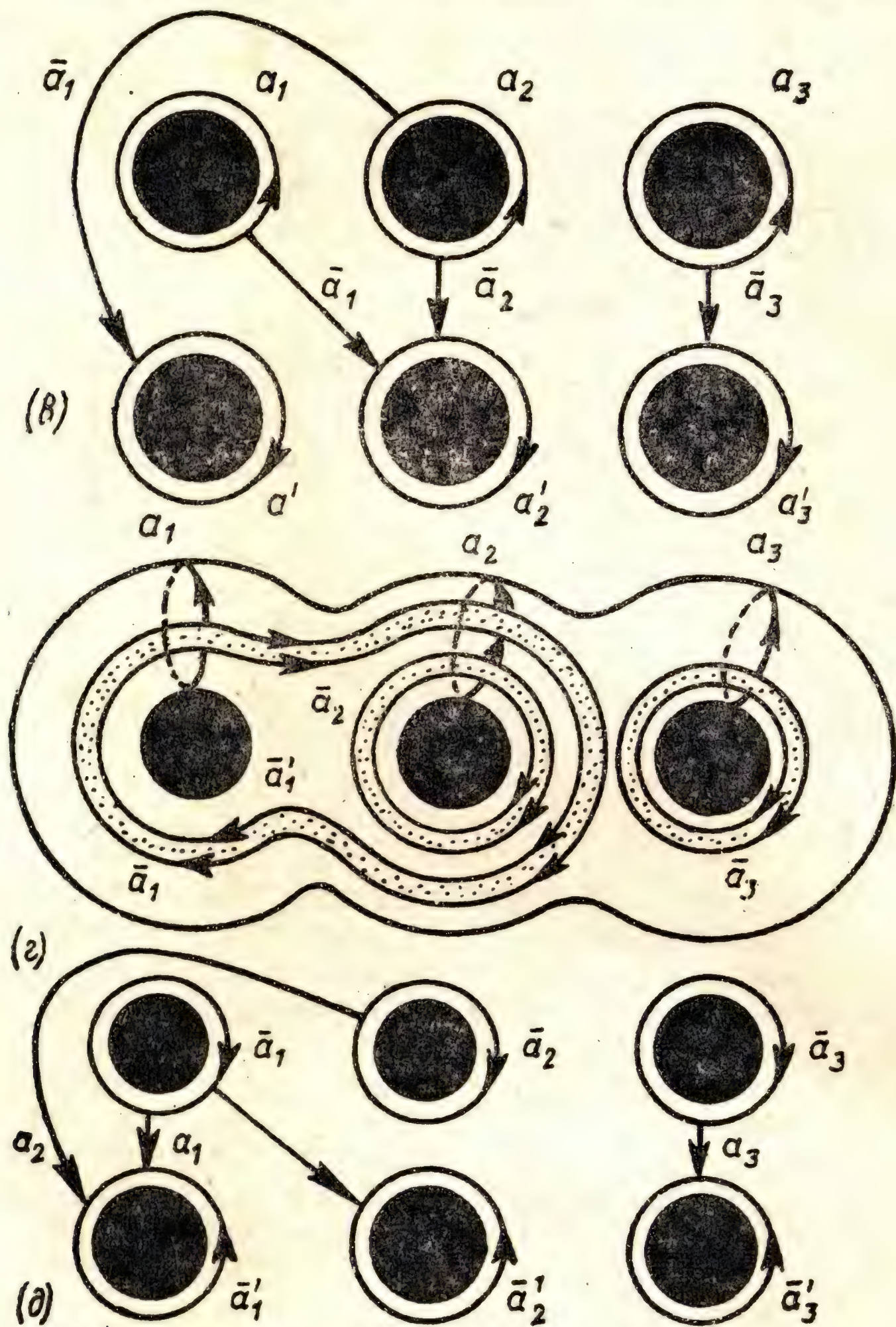


Рис. 7 (в) — (д)

ружности индекса два также оказываются ориентированными. При разрезании кренделя по всем окружностям индекса один, окружности индекса два разрываются на части и превращаются в набор ориентированных отрезков, соединяющих точки, расположенные на границах разрезов. Снабдим каждый из отрезков ориентацией и номером, индуцированными на нем

окружностью индекса два, частью которой он является. Тем самым оказывается, что мы получаем на двумерной сфере с $2g$ дырками некоторый набор гладких дуг, соединяющих точки, расположенные на границах дырок. Пример см. на рис. 7. Изображен крендель рода 3 и два графа, получающиеся при его разрезании вдоль окружностей индекса один (рис. 7(в)) и вдоль окружностей индекса два (рис. 7(д)). Для удобства изображения этих графов из сферы с $2g$ дырками выколота точка, не принадлежащая графам, после чего получившийся диск с $2g$ дырками сделан плоским. В дальнейшем мы будем изображать получающиеся графы на обычной плоскости, из которой вырезано $2g$ непересекающихся дисков.

Определение 2. Пусть α — произвольный диффеоморфизм кренделя рода g . Рассмотрим соответствующую ему сеть s . Графом Уайтхеда $W_i(s)$ индекса i назовем граф, получающийся на плоскости в результате описанного выше разреза кренделя по окружностям индекса i . Таким образом, каждый диффеоморфизм α определяет два графа Уайтхеда $W_1(s)$ и $W_2(s)$, где $s = s(\alpha)$.

Эти два графа, вообще говоря, не совпадают. Вершинами графа Уайтхеда являются окружности разрезов, ребрами — части окружностей другого индекса, разорванные разрезами в набор отрезков. При этом мы считаем, что все ребра и окружности ориентированы и снабжены соответствующими индексами. Ясно, что задание любого из этих обоих графов полностью определяет исходную сеть. Для этого достаточно склеить соответствующие пары окружностей и восстановить крендель. Для простоты мы будем иногда изображать вершины графов Уайтхеда не окружностями (дырками), а точками. Отметим, что диффеоморфизм α восстанавливается по сети $s = (a_i, \bar{a}_j)$ неоднозначно. Впрочем, это обстоятельство для нас несущественно, так как все диффеоморфизмы, определяющие одну и ту же сеть, определяют одно и то же многообразие.

Итак, можно считать, что пары графов Уайтхеда кодируют множество всех трехмерных многообразий. Каждое многообразие представлено бесконечным множеством графов Уайтхеда. Рассмотрим простейшее трехмерное многообразие — стандартную сферу. Ей отвечает некоторое множество графов Уайтхеда, содержащееся в множестве всех графов Уайтхеда всех трехмерных многообразий. Еще А. Пуанкаре в начале века поставил вопрос: как распознать стандартную сферу в множестве всех трехмерных многообразий, заданных каким-либо списком? Подробности см., например, в [34]. С этой задачей тесно связана знаменитая проблема Пуанкаре: верно ли, что любая трехмерная гомотопическая сфера гомеоморфна стандартной? Напомним, что многообразие называется гомотопической сферой, если имеет те же гомотопические группы, что и сфера.

В настоящем параграфе мы сосредоточим основное внимание на задаче алгоритмического распознавания трехмерной

сферы в классе диаграмм Хегора рода два. Другими словами, требуется найти алгоритмически распознаваемое (и легко вычислимое) характеристическое свойство тех графов Уайтхеда, которые отвечают стандартной сфере. Если такое свойство удастся найти, это и позволит отвечать на вопрос: является ли произвольно заданная пара графов Уайтхеда графами, отвечающими стандартной сфере, или же они описывают другое многообразие, не гомеоморфное сфере? Мы не будем анализировать здесь тонкости понятия алгоритмического распознавания, ограничившись интуитивным представлением, согласно которому мы должны «сконструировать вычислительную машину», действующую по какой-то универсальной единой программе и дающую ответ на вопрос: является ли код многообразия, поданный на «вход машины», кодом стандартной сферы или нет. Различные подходы к решению этого вопроса см. в [40—45].

Для дальнейшего удобно рассматривать сеть с точностью до произвольного диффеоморфизма кренделя на себя. Другими словами, окружности a_i не обязаны быть теперь стандартными меридианами, а могут быть их образами при произвольном диффеоморфизме. Если две сети эквивалентны в этом смысле, т. е. переводятся друг в друга некоторым диффеоморфизмом, то легко проверить, что соответствующие им многообразия диффеоморфны. Поэтому, заменяя сеть на её эквивалентную, мы не меняем топологический тип многообразия, но зато избавляемся от необходимости постоянно помнить о том, что окружности a_i обязательно изображаются стандартными меридианами. В частности, мы устраняем некоторое неравноправие между окружностями индексов один и два, возникшее в тот момент, когда в основу определения сети мы положили окружности a_i . Если это нам потребуется, то, применяя к сети подходящий диффеоморфизм, мы всегда можем добиться того, чтобы окружности индекса один превратились в стандартные меридианы. Аналогичным образом можно превратить в стандартные параллели окружности индекса два. При этом, конечно, окружности индекса один отнюдь не отобразятся в стандартные меридианы, а превратятся в какие-то, быть может сложные, окружности на кренделе.

Приведем пример некоторой простейшей сети s_0 , определяющей стандартную сферу. Возьмем в качестве a_i стандартные меридианы, а в качестве $\bar{a}_j$ — стандартные параллели. Диффеоморфизм кренделя, задающий такую сеть (определенный неоднозначно), переставляет параллели и меридианы. Склеивая два заполненных кренделя по такому диффеоморфизму и используя информацию § 1 настоящей главы, мы и получаем стандартную сферу. Соответствующие этой сети графы Уайтхеда показаны на рис. 8. Они являются простейшими в множестве всех графов Уайтхеда. Оказывается, любая сеть s стандартной сферы может быть получена из простейшей сети s_0

(того же рода) применением к сети s_0 некоторого числа простых операций, которые мы сейчас опишем.

Пусть s — некоторая сеть какого-либо многообразия M , не обязательно сферы. Рассмотрим на кренделе две окружности индекса один a_i и a_j . Пусть γ — некоторый гладкий путь, соединяющий точку окружности a_j с некоторой точкой окружности a_i и не имеющий других точек пересечения с окружностями

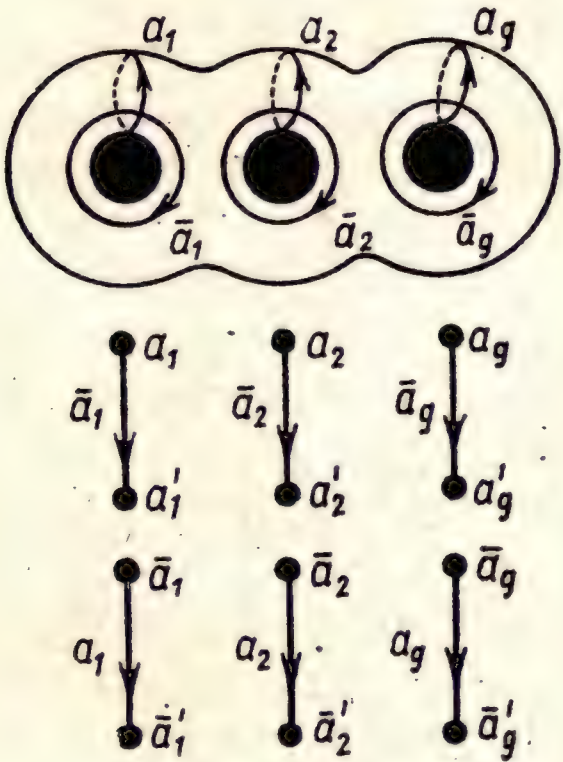


Рис. 8

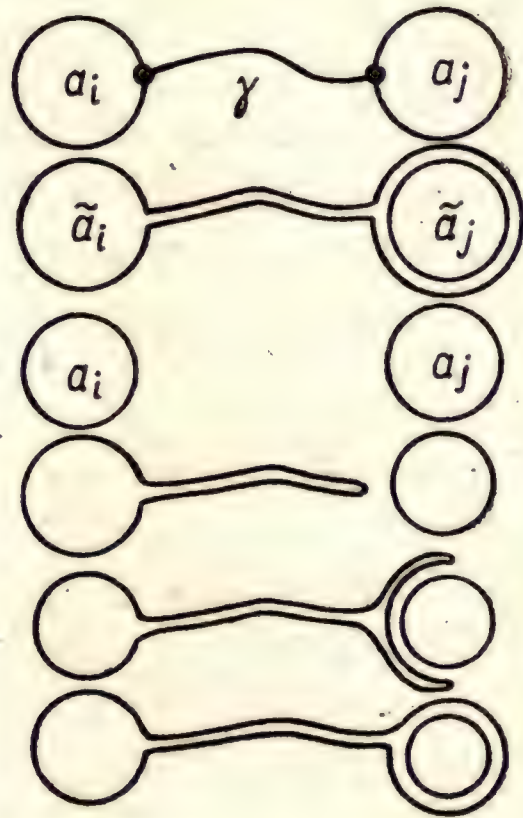


Рис. 9

индекса один. Заменяем пару окружностей a_i, a_j на новую пару окружностей $\tilde{a}_i, \tilde{a}_j$, где $\tilde{a}_j = a_j$, а новая окружность $\tilde{a}_i$ изображена на рис. 9. Другими словами, из окружности a_i вырастает тонкий язык, который скользит вдоль пути γ , затем, подходя к окружности a_j , он раздваивается и начинает охватывать окружность a_j . Наконец, происходит «заглатывание» окружности a_j , и окончательный результат показан на рис. 9. Окружность a_j при этом не меняется.

Определение 3. Описанная выше операция называется операцией индекса один. Аналогичным образом определяется операция индекса два.

Отметим, что отрезок γ из определения операции индекса один мог пересекать окружности индекса два. Как всегда, мы считаем все пересечения трансверсальными, т. е. находящимися в общем положении.

Лемма 1. Пусть к некоторой сети s , определяющей некоторое многообразие M , применены в произвольном числе и порядке операции индексов один и два. Тогда в результате мы получаем некоторую сеть (в смысле определения 1) того же рода, что и исходная сеть. Далее, эта новая сеть определяет многообразие, диффеоморфное исходному. Таким образом, операции индексов один и два не меняют топологического типа трехмерного многообразия.

Доказательство леммы можно провести методами теории Морса и изложено, например, в [40]. Таким образом, мы получаем возможность, исходя из какой-то фиксированной сети, строить бесконечное множество сетей, задающих то же самое многообразие. При этом операции индексов один и два могут применяться в произвольных комбинациях. Эти операции, кстати, в общем случае не позволяют получить все сети данного многообразия (и данного рода), отправляясь от какой-то одной сети. Но в одном случае это удастся (Вальдхаузен).

Предложение 1. Пусть s_0 — простейшая сеть рода g стандартной сферы, показанная на рис. 8. Тогда любая другая сеть стандартной сферы того же рода может быть получена из сети s_0 в результате применения некоторого числа операций индексов один и два.

Доказательство см., например, в [43, 40]. Таким образом, множество сетей, получающихся из простейшей сети применением операций индексов один и два, в точности совпадает с множеством всех сетей, соответствующих стандартной сфере. Важным обстоятельством является также то, что любая сеть некоторого фиксированного рода может быть получена операциями указанного типа из простейшей сети того же рода. Отметим, что применение операций индексов один и два не меняет род сети. Предложение 1 дает описание всех кодов стандартной сферы в пространстве всех кодов всех трехмерных многообразий. Не следует, впрочем, думать, что это дает алгоритм распознавания кодов сферы. Дело в том, что в дополнение к сказанному мы должны теперь научиться распознавать те сети, которые получаются из простейшей применением описанных операций.

2. Волны, разбивающие вершины, и операция упрощения графов Уайтхеда.

Рассмотрим некоторую сеть на кренделе какого-то рода. Напомним, что мы рассматриваем только ориентированные сети (пункт 1). Сеть задает разбиение кренделя на некоторые области U .

Определение 4. Область U назовем отмеченной, если среди ребер, задающих ее границу, найдутся два ребра β_1 и β_2 , принадлежащие к одной окружности, и их ориентации согласованы с одним из обходов границы области. Ребра β_1 и β_2 назовем отмеченными (рис. 10). Волной τ назовем отрезок внутри отмеченной области, соединяющий внутренние точки отмеченных ребер β_1 и β_2 (рис. 10).

Свойства волны подробно изучены в [40]. Это понятие оказалось полезным для построения алгоритма распознавания диаграмм Хегора сферы (для рода два). Переформулируем определение волны на языке графов Уайтхеда. Оказывается, наличие волны в сети особенно наглядно проявляется именно на этих графах.

Пусть, для определенности, отмеченные ребра (см. опре-

деление 4) принадлежат окружности a_i индекса один. Разрежем крендель по окружностям индекса один, в результате чего получим граф $W_1(s)$. Окружность a_i превратится при этом в две вершины a_i и a_i' графа $W_1(s)$. Волна, т. е. отрезок τ , превращается в гладкий отрезок, начинающийся и кончающийся в одной из этих вершин. В какой именно — зависит от соотношения ориентаций. Пусть, для определенности, отрезок τ начинается и кончается в a_i . Замечательным свойством дуги τ является то, что она пересекает

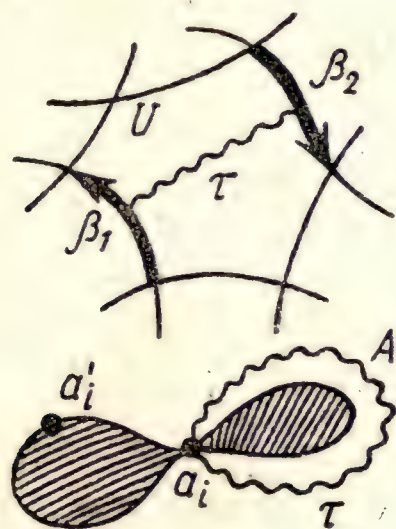


Рис. 10

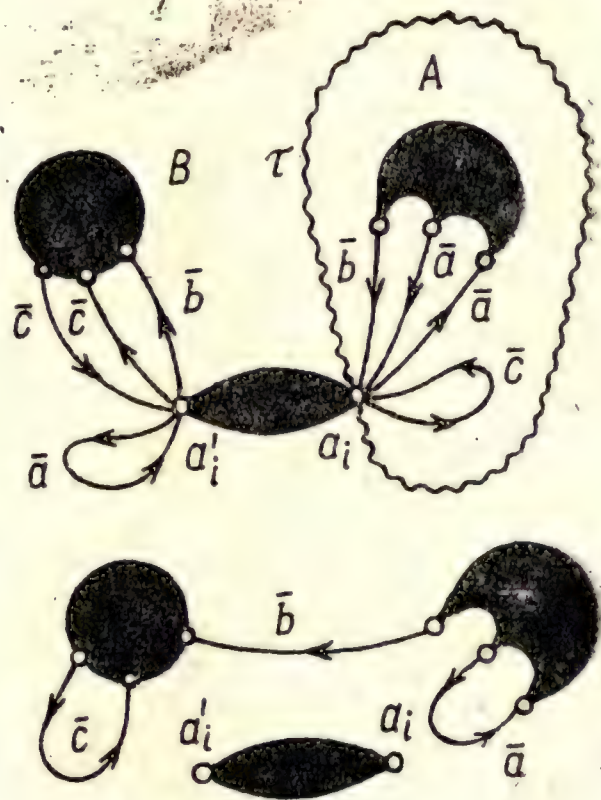


Рис. 11

ся с графом $W_1(s)$ только в точке a_i (рис. 10). Следовательно, выбросив вершину a_i из графа, мы разбиваем его по крайней мере на две связные непустые компоненты. Одна из них расположена в области A , ограниченной на плоскости волной τ , а другая расположена вне этой области (рис. 10). Будем считать, что $a_i' \notin A$ и что a_i лежит хотя бы на одном ребре из B .

Определение 5. Вершину графа Уайтхеда, в которой начинается и заканчивается волна τ , назовем разбивающей вершиной, если существует хотя бы одно ребро графа, расположенное в A и содержащее вершину a_i .

Такие вершины впервые появились в замечательной работе Уайтхеда [42], в которой была решена задача распознавания наборов элементов в свободной группе, являющихся ее образующими (так называемых простых наборов).

Роль разбивающей вершины заключается в том, что ее наличие на одном из графов Уайтхеда позволяет упростить оба графа Уайтхеда и, следовательно, диаграмму Хегора многообразия. Эквивалентным образом, наличие волны в сети s позволяет упростить сеть и перейти при помощи операций индексов один и два к новой сети, имеющей меньшее число ребер. Опишем процедуру упрощения графа Уайтхеда.

Пусть a_i — разбивающая вершина на графе $W_1(s)$. Волна τ разбивает плоскость на две области: A и B , одна из которых содержит вершину a_i' . Напомним, что каждая окружность a_i порождает две вершины a_i и a_i' при разрезании кренделя.

Пусть A — область, не содержащая a_i' (рис. 11). Согласно определению разбивающей вершины существует хотя бы одно ребро графа, расположенное в A и содержащее вершину a_i . Будем говорить, что ребро графа Уайтхеда инцидентно с некоторой вершиной, если ребро входит в вершину или выходит из нее. Рассмотрим множество всех ребер графа Уайтхеда, инцидентных с вершиной a_i и расположенных в области A , окруженной волной (рис. 11). Эти ребра разбиваются на три класса.

1) Ребра первого класса характеризуются тем, что они являются частью окружности (индекса два), следующим образом взаимодействующей с вершиной a_i . Эта часть окружности, изображенная на рис. 11 буквой $\bar{a}$, входит в вершину a_i , затем «перескакивает» в вершину a_i' , выходит из нее, делает петлю и возвращается снова в вершину a_i' , не проходя ни через какие другие вершины графа. Далее, вернувшись в a_i' , часть окружности $\bar{a}$ «перескакивает» обратно в вершину a_i и, выходя из нее, возвращается в часть графа Уайтхеда, окруженную волной. 2) Ребра второго класса характеризуются тем, что они являются частями окружности, следующим образом взаимодействующей с вершиной a_i . Эта часть окружности, изображенная на рис. 11 буквой $\bar{b}$, входит в вершину a_i , «перескакивает» в вершину a_i' , а затем входит в некоторую другую вершину графа, отличную от вершины a_i' и расположенную вне области A . 3) Ребра третьего класса характеризуются тем, что они являются частями окружности, следующим образом взаимодействующей с вершиной a_i . Эта часть окружности, изображенная на рис. 11 буквой $\bar{c}$, выходит из некоторой вершины графа, расположенной вне области A , входит в вершину a_i' , «перескакивает» в вершину a_i , затем делает петлю и возвращается снова в вершину a_i , не проходя ни через какие другие вершины графа. После этого часть окружности $\bar{c}$ «перескакивает» обратно в вершину a_i' и, выйдя из нее, входит в некоторую вершину графа, расположенную вне области A .

Опишем перестройку, операцию упрощения графа Уайтхеда при условии наличия на нем хотя бы одной разбивающей вершины a_i . Эта операция следующим образом преобразует граф Уайтхеда. 1) В случае ребер первого класса аннулируется петля $\bar{a}$, начинающаяся и кончающаяся в вершине a_i' , а два ребра, инцидентные с вершиной a_i , аннулируются и заменяются одним ребром, соединяющим те же вершины графа, отличные от a_i , с которыми были первоначально инцидентны эти ребра (рис. 11). 2) В случае ребер второго класса оба ребра $\bar{b}$, инцидентные с вершинами a_i и a_i' соответственно, аннулируются и заменяются одним ребром, которое, минуя вершины a_i и a_i' , напрямик соединяет те же вершины графа, отличные от a_i и a_i' , с которыми были первоначально инцидентны эти ребра (рис. 11). 3) В случае ребер третьего класса операция совпадает с операцией, описанной выше для случая ребер первого

класса. Единственное отличие заключается в том, что нужно поменять ролями вершины a_i и a_i' (рис. 11).

Операция перестройки графа Уайтхеда полностью описана. Однако она нуждается в комментариях. Дело в том, что операция была описана абстрактно, без учета, например, того обстоятельства, что а priori возможность реализации описанных перестроек в плоскости отнюдь не очевидна.

Теорема 1 (см. [40]). а) Пусть $\tilde{W}$ — граф, полученный из графа Уайтхеда $W_i(s)$ описанной выше перестройкой при наличии хотя бы одной разбивающей вершины. Тогда граф $\tilde{W}$ является плоским графом, т. е. допускает реализацию на плоскости, и, более того, все описанные выше перестройки могли быть с самого начала реализованы в плоскости. б) Граф $\tilde{W}$ является графом Уайтхеда для некоторой новой сети $\tilde{s}$, т. е. $\tilde{W} = W_i(\tilde{s})$. в) Новая сеть $\tilde{s}$ получается из исходной сети s (которой соответствовал исходный граф Уайтхеда) путем применения операций индексов один и два (в действительности достаточно операций одного индекса). г) Сеть $\tilde{s}$ задает то же самое многообразие (с точностью до диффеоморфизма), что и исходная сеть s , т. е. описанная операция перестройки сохраняет топологический тип многообразия. д) Новый граф Уайтхеда $W_i(\tilde{s})$ имеет меньше ребер, чем исходный граф (при том же количестве вершин), а потому является более простым.

Мы называли операцию перестройки операцией упрощения графа Уайтхеда именно в силу утверждения д) теоремы 1. Новый граф $W_i(\tilde{s})$ проще исходного в том смысле, что число его ребер строго меньше числа ребер у первоначального графа. Утверждение д) очевидно из рис. 11, из которого видно, что операция перестройки уменьшает число ребер по крайней мере на единицу. Утверждение г) очевидно следует из утверждения в) и из леммы 1. Более сложными являются утверждения б) и в), см. [40]. Таким образом, если сеть s некоторого многообразия M^3 содержит волну (т. е. хотя бы на одном из ее графов Уайтхеда $W_i(s)$ есть разбивающая вершина), то существует чрезвычайно простой алгоритм, который позволяет перевести сеть s в новую сеть $\tilde{s}$, задающую то же многообразие и являющуюся более простой, чем исходная.

3. Существование разбивающей вершины на графах Уайтхеда рода два, соответствующих стандартной трехмерной сфере. Алгоритм распознавания сферы.

Простейшая сеть s_0 , соответствующая стандартной сфере, не содержит ни одной разбивающей вершины (рис. 8). Однако сеть, показанная на рис. 7, соответствует стандартной сфере (проверьте!), и оба ее графа Уайтхеда содержат разбивающие вершины (рис. 8). Более того, можно доказать, что если сеть s получена из простейшей сети s_0 операциями одного индекса, то s обязательно содержит по крайней мере одну волну (при условии, что $s \neq s_0$). При этом следует иметь в виду, что разби-

вающая вершина может иметься только на одном из двух графов Уайтхеда. Простой пример сети, соответствующей стандартной сфере, показан на рис. 12. Здесь же изображены два графа Уайтхеда. Один из них содержит две разбивающие вершины, другой — ни одной (рис. 12(б)).

В результате анализа фундаментальной работы Уайтхеда [42] И. А. Володин и А. Т. Фоменко сформулировали в 1974 г.

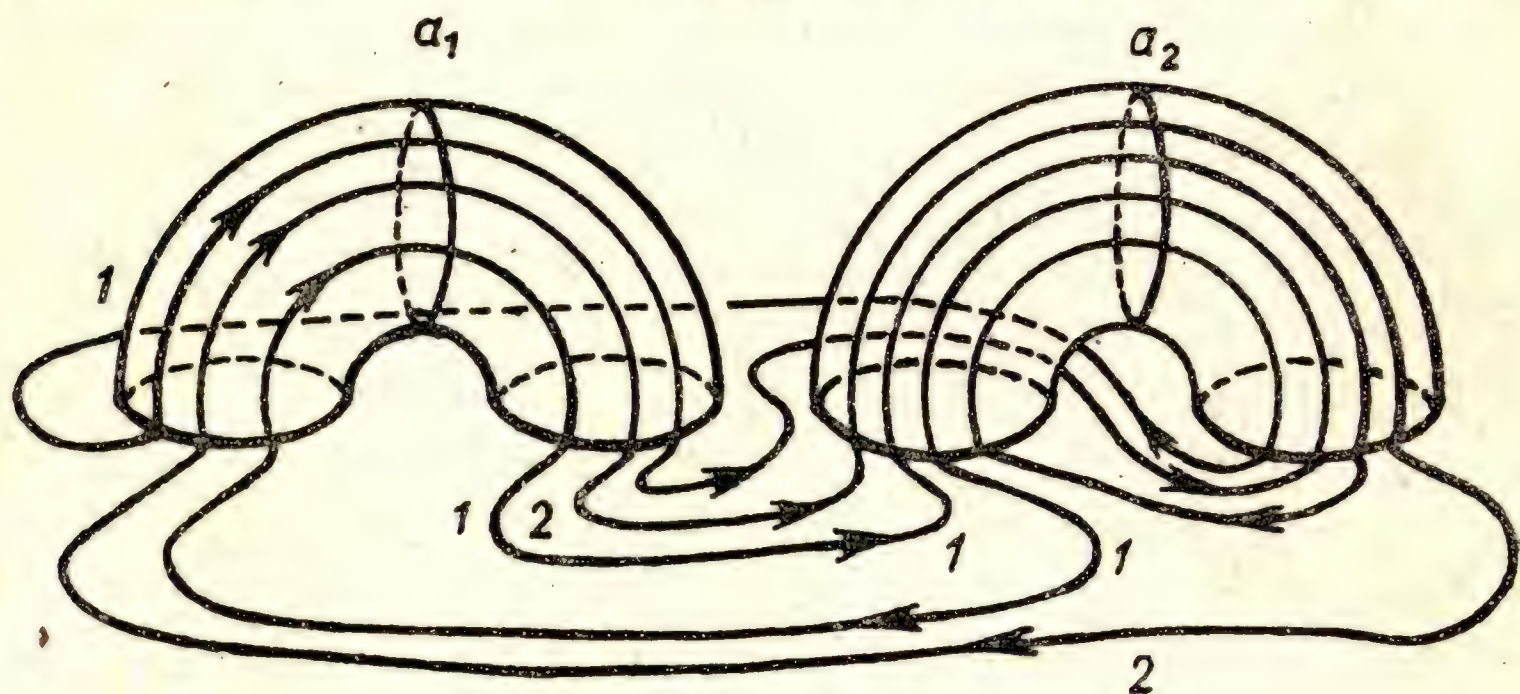


Рис. 12(а)

следующую гипотезу: любая сеть, соответствующая стандартной сфере и отличная от простейшей, содержит по крайней мере одну волну [40, 48]. Для проверки гипотезы была использована ЭВМ БЭСМ-6, на которой моделировался процесс построения сетей с помощью операций индексов один и два [40]. Оказалось, что все 10^6 сетей, построенных ЭВМ случайным образом, и задающих стандартную сферу, действительно содержат по крайней мере одну волну. Однако затем теоретическим путем было обнаружено (см. [47]), что для сетей рода три и выше сформулированная гипотеза все-таки неверна (см. также [49, 50]). Таким образом, ситуация оставалась невыясненной для сетей рода два. (В случае сетей рода один ответ тривиален, см. § 1 гл. 3.) Вследствие этого вычислительный эксперимент был повторен автором настоящей книги уже специально для сетей рода два с учетом в программе для ЭВМ обнаруженных геометрических механизмов, приводящих иногда к уничтожению волны на сетях сферы рода три и выше.

Оказалось, что, несмотря на ограниченность возможностей ЭВМ, этот расширенный эксперимент не обнаружил ни одного контрпримера к гипотезе о существовании волны на сетях сферы рода два. И наконец, в 1980 г. в работе [46] эта гипотеза Володина и Фоменко была окончательно теоретически доказана для всех сетей сферы рода два.

Теорема 2 (см. [46]). Каждая сеть рода два стандартной сферы, отличная от простейшей, содержит по крайней мере одну волну. В частности, хотя бы один из графов Уайтхеда содержит по крайней мере одну разбивающую вершину.

Доказательство этого факта весьма нетривиально и поэтому здесь опущено. Отметим любопытный результат, попутно полученный в [46] и описывающий структуру произвольного графа Уайтхеда рода два, отвечающего трехмерному многообразию. Оказывается, каждый граф Уайтхеда изоморфен одному из плоских графов типа I, II, III, показанных на рис. 13, где че-

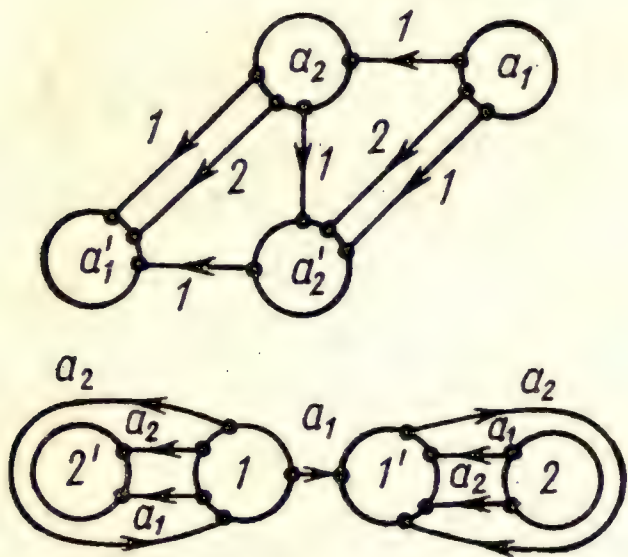


Рис. 12(б)

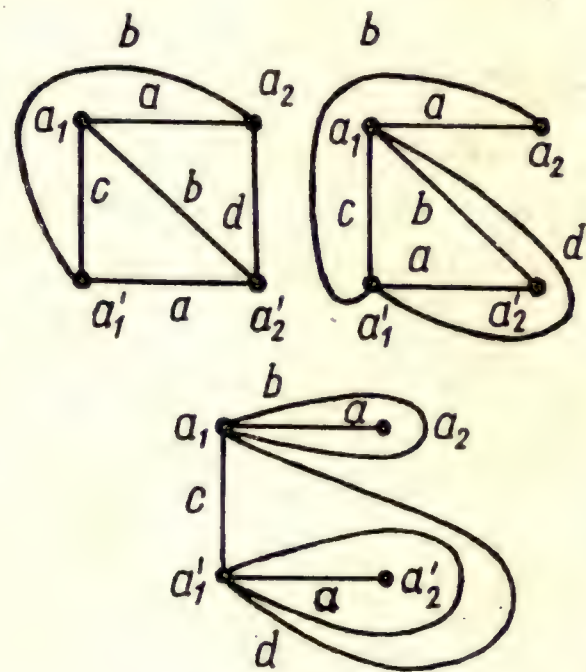


Рис. 13

рез a, b, c, d обозначены числа «параллельных ребер», соединяющих соответствующие пары вершин графа.

Таким образом, подтверждена и окончательно обоснована эффективность следующего чрезвычайно простого и наглядного алгоритма распознавания стандартной сферы в классе диаграмм Хегора рода два, сформулированного Володиным и Фоменко и первоначально проверенного на ЭВМ [40]. А именно, берем произвольную сеть s и выясняем, содержит ли она хотя бы одну волну, т. е. содержит ли хотя бы один из ее графов Уайтхеда разбивающую вершину. Если разбивающих вершин нет на обоих графах, то исходная сеть задает трехмерное многообразие, не диффеоморфное стандартной сфере, и мы получаем ответ на первом же шаге. Если же имеется хотя бы одна разбивающая, то мы применяем операцию упрощения сети (графов Уайтхеда), носящую простой алгоритмический характер. В результате получаем новую сеть s' , соответствующую тому же многообразию. Снова выясняем: имеется ли разбивающая вершина. Если разбивающей нет, то исходное многообразие не является сферой. Если разбивающая есть, то упрощаем сеть и т. д. Таким образом, в результате конечного числа шагов, не превышающего числа ребер у исходной сети, мы либо остановимся на сети, не имеющей ни одной волны (в этом случае исходное многообразие не является сферой), либо остановимся на простейшей сети s_0 , соответствующей каноническому заданию сферы. В этом последнем случае исходное многообразие диффеоморфно сфере. Этот алгоритм дополняет известный алгоритм Хакена.

§ 3. О решении четырехмерной проблемы Пуанкаре: любая гомотопическая четырехмерная сфера гомеоморфна стандартной сфере

В 1904 году в своем труде «Analysis situs» французский математик А. Пуанкаре высказал гипотезу: если трехмерное многообразие M имеет такие же целочисленные одномерные гомологии, что и стандартная трехмерная сфера S^3 , т. е. $H_1(M, \mathbb{Z}) = 0$, то M гомеоморфно сфере S^3 . Это предположение оказалось неверным. Сам Пуанкаре построил трехмерное многообразие M_0 с нулевой одномерной группой гомологий, но не гомеоморфное сфере. Дело в том, что это многообразие оказалось неодносвязным. Его фундаментальная группа имеет вид $\pi_1(M_0) = \langle a, b \mid (ab)^2 = a^3 = b^5 \rangle$. Трехмерные многообразия, у которых $H_1(M, \mathbb{Z}) = 0$, называются гомологическими трехмерными сферами. Многообразие, построенное Пуанкаре, называется иногда пространством додекаэдра. Подробности см., например, в [3, 34, 55].

Ввиду указанного обстоятельства гипотеза приняла следующий окончательный вид: если $\pi_1(M^3) = 0$, то многообразие M^3 гомеоморфно сфере S^3 . Из двойственности Пуанкаре легко следует, что если $\pi_1(M^3) = 0$, то M^3 гомотопически эквивалентно S^3 . Поэтому гипотезу Пуанкаре можно сформулировать так: если $M^3 \sim S^3$ (знаком $\sim$ мы обозначаем гомотопическую эквивалентность), то $M^3 = S^3$. В этом виде гипотеза имеет смысл для многообразий любых размерностей, а именно если $M^n \sim S^n$, то $M^n = S^n$. Отметим, что для многообразий размерности 4 и выше односвязность многообразия отнюдь не означает, что оно гомотопически эквивалентно сфере.

На самом деле в формулировке гипотезы Пуанкаре необходима точность. Надо оговаривать, какой класс многообразий мы имеем в виду и в каком смысле мы понимаем равенство $M^n = S^n$. В настоящее время в топологии многообразий наиболее употребительны три категории: Тор — топологических многообразий и гомеоморфизмов; PL — кусочно-линейных многообразий (т. е. многообразий, правильным образом разбитых на симплексы так, что эта триангуляция локально устроена как прямолинейная триангуляция евклидова пространства, см. гл. 1) и кусочно-линейных гомеоморфизмов (т. е. гомеоморфизмов, сохраняющих некоторые подразделения триангуляции, см. гл. 1), и, наконец, Diff — гладких многообразий и диффеоморфизмов. Таким образом, в каждой размерности n имеет смысл рассматривать три возможных варианта гипотезы Пуанкаре $P_n(\text{CAT})$, где CAT условно обозначает категории Тор, PL или Diff.

Итак, гипотеза $P_n(\text{CAT})$ формулируется так: если M^n есть CAT-многообразие и $M^n \sim S^n$, то $M^n = S^n$, т. е. M^n эквивалентно (гомеоморфно, кусочно-линейно гомеоморфно, диффеоморфно) сфере S^n в смысле CAT.

Приведем сводку известных результатов до 1981 года, связанных с гипотезой Пуанкаре. Эта сводка будет краткой и отнюдь не претендует на исчерпывающую полноту.

а) Имеются соотношения между гипотезами $P_n(\text{CAT})$.

1. Если $n \neq 4$, то $P_n(\text{Top}) = P_n(\text{PL})$. См. работы Kirby—Siebenmann (1970 г.) для $n \geq 5$, Mois (1952 г.) для $n = 3$, Pao (1925 г.) для $n = 2$. При $n = 1$ утверждение тривиально.

2. Если $n < 7$, то $P_n(\text{Diff}) = P_n(\text{PL})$.

б) Укажем размерности, для которых гипотеза Пуанкаре справедлива.

1. Гипотеза $P_n(\text{PL})$ верна при $n \neq 3, 4$. См. работы Смейла (1961 г.) при $n \geq 5$. При $n = 2$ справедливость гипотезы следует из теоремы классификации двумерных поверхностей.

2. Гипотеза $P_n(\text{Diff})$, вообще говоря, не верна при $n \geq 7$. Милнор (1956 г.) доказал существование нестандартных гладкостей на сферах при $n \geq 7$.

Вся эта информация может быть следующим образом условно изображена в виде схемы на рис. 14.

В конце 1981 года американский математик М. Фридман доказал справедливость гипотезы $P_4(\text{Top})$, т. е. поставил «+» в кружок около четверки на рис. 14. Он доказал, что замкнутое топологическое четырехмерное многообразие M^4 , гомотопически эквивалентное сфере S^4 , гомеоморфно ей. На самом деле он доказал теорему об h -кобордизме (см. [3]) для четырехмерных многообразий. Гипотеза Пуанкаре легко вытекает из этой теоремы [56, 62].

Напомним, что компактное $(n+1)$ -мерное PL-многообразие W^{n+1} , имеющее край, состоящий из двух компонент: M_1^n и M_2^n , называется h -кобордизмом, если пленка W деформационно ретрагируется на оба края M_i^n , $i = 1, 2$, и если $\pi_1(W) = 0$. Простейшим примером h -кобордизма служит прямое произведение

$M_1 \times I$. Теорема Смейла утверждает, что если $n \geq 5$, то $W = M_1 \times I$, где W h -кобордизм. В частности, отсюда следует, что края M_1 и M_2 h -кобордизма PL-гомеоморфны.

PL-гипотезу Пуанкаре для $n \geq 5$ из теоремы Смейла можно вывести так. Пусть $M^n \sim S^n$. Удалим из многообразия малый открытый шар. В результате получим n -мерное многообразие V^n . Рассмотрим $V^n \times I$ и удалим шар из этого прямого произведения. Мы получили h -кобордизм между двумя многообразиями: S^n и $\partial(V^n \times I)$. По теореме Смейла этот h -кобордизм три-

виален, и, следовательно $S^n \stackrel{\text{PL}}{=} \partial(V^n \times I)$. Остается заметить, что V^n лежит в $\partial(V^n \times I) = S^n$ и ограничивается там сферой, S^{n-1} . По теореме Шенфлиса имеем: $V^n \stackrel{\text{Top}}{=} \text{стандартному } n\text{-мерно-}$

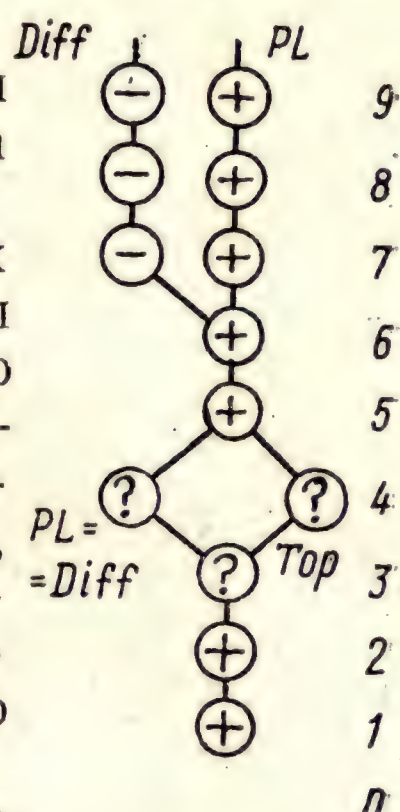


Рис. 14

му шару. Итак, $M^n \stackrel{\text{Top}}{=} S^n$, что и требовалось (рис. 15). Приведенное выше рассуждение можно уточнить, и получить, что $M^n \stackrel{\text{PL}}{=} S^n$.

Фридман доказал, что теорема об h -кобордизме справедлива и для $n=4$, но только в несколько более слабом смысле, а именно если W^5 является PL- h -кобордизмом, то $W^5 \stackrel{\text{Top}}{=} M_1 \times I$. Здесь многообразие W^5 предполагается компактным. Однако в действительности эта теорема справедлива и для некомпактного h -кобордизма W^5 между некомпактными многообразиями. Для этого нужно дополнительно потребовать, чтобы непрерывные деформации пленки W^5 на ее края M_i^4 были собственными (т. е. прообраз компакта должен быть компактом) и чтобы пленка W^5 была «односвязна в бесконечности». Последнее ус-

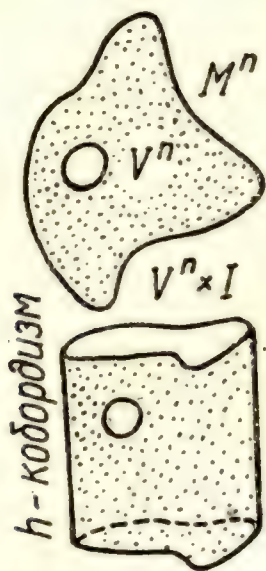


Рис. 15

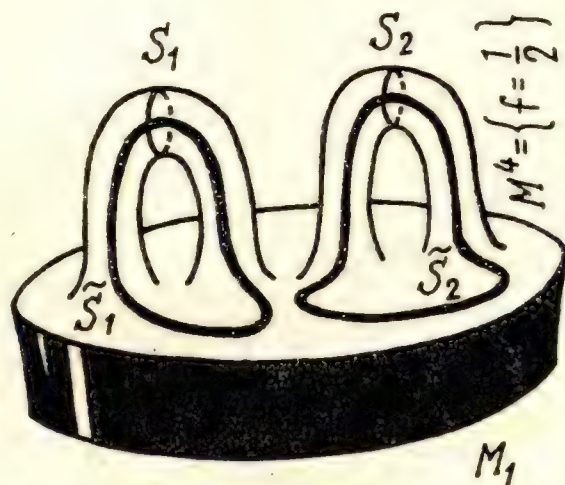


Рис. 16

ловие означает, что «далекие» петли стягиваются в точку, оставаясь «далекими», т. е. стягивающая гомотопия не должна затрагивать заранее заданного компакта, сколь угодно большого. Гипотеза $P_4(\text{Top})$ легко вытекает из этой теоремы о некомпактном h -кобордизме. Если $M^4 \sim S^4$ является PL-многообразием, то уже приведенное выше рассуждение полностью повторяется. Если же M^4 не является PL-многообразием, то $M^4 \setminus *$ стягиваемо (здесь через $*$ обозначена точка), следовательно параллелизуемо. Из теории погружений вытекает, что $M^4 \setminus *$ погружается в евклидово пространство $\mathbb{R}^4$ и, следовательно, имеет PL- (или Diff-) структуру. Дословно повторяя вышеприведенное рассуждение, мы получаем h -кобордизм W^5 между S^4 и $\partial(V \times I)$, являющийся PL-многообразием вне отрезка или, что то же самое, вне точки. Теперь можно выбросить отрезок, проходящий через эту точку и соединяющий края кобордизма, после чего нужно применить теорему о некомпактном h -кобордизме к оставшемуся многообразию.

Приведем краткую схему доказательства теоремы об h -кобордизме между 4-мерными многообразиями. При этом мы будем опираться на сведения, содержащиеся в [3, 62].

1. Доказательство Смейла теоремы об h -кобордизме в размерностях $n \geq 5$ проводится в размерности 4 до тех пор, пока не возникает необходимости применить так называемый прием Уитни.

2. Детальный анализ приема Уитни выявляет 4 причины, по которым этот прием не действует в размерности 4.

3. Так называемая редукция Кассона позволяет с помощью обычных финитных методов топологии многообразий справиться с тремя из этих трудностей и с помощью некоторого бесконечного процесса частично справиться с четвертой трудностью.

4. Теорема Фридмана о гомеоморфности ручки Кассона настоящей ручки позволяет преодолеть последние трудности. Опишем вкратце последовательные шаги в реализации этой программы.

Шаг первый. Доказательство Смейла теоремы об h -кобордизме таково. Сначала h -кобордизм W^{n+1} разбивается на ручки. Предварительно выделяется малая окрестность краев. Потом с помощью специальных операций с ручками это первоначальное разбиение мы стараемся упростить. Конечная цель — устранить лишние ручки. Этими приемами удастся убрать все ручки индексов $0, 1, \dots, k-1$ (здесь мы считаем, что $n=2k$), а также, в силу двойственности удастся убрать все ручки индексов $n+1, n, \dots, k+2$. В результате остаются только ручки индексов k и $k+1$. Применяя это рассуждение в случае $n=4$, мы получаем, что остаются лишь ручки индексов 2 и 3. Более того, удастся добиться, чтобы ручки индексов k и $k+1$ разбивались на алгебраически дополнительные пары. Это означает следующее. В многообразии M^4 , разделяющем ручки индексов k и $k+1$, содержатся два набора вложенных сфер. Первый набор $S = \{S_1^k, S_2^k, \dots, S_m^k\}$ состоит из правосторонних сфер (границ сепаратрисных дисков размерности $k+1$), вложенных в поверхность уровня $M^4 = \{f=1/2\}$ для соответствующей гладкой функции Морса. Эти правосторонние сферы соответствуют ручкам индекса k . Вторым набором $\tilde{S} = \{\tilde{S}_1^k, \tilde{S}_2^k, \dots, \tilde{S}_m^k\}$ состоит из левосторонних сфер, вложенных в то же самое многообразие $M^4 = \{f=1/2\}$. Эти сферы являются границами сепаратрисных дисков размерности $k+1$, выходящих из критических точек индексов $k+1$. Эти точки соответствуют ручкам индекса $k+1$. Можно считать, что левосторонние и правосторонние сферы пересекаются трансверсально, в частности, в конечном числе точек. Алгебраическая дополнительность сфер означает, что индекс пересечения μ этих сфер имеет вид $\mu(S_i^k, \tilde{S}_j^k) = \delta_{ij}$ (рис. 16). Если бы мы смогли алгебраическую дополнительность этих двух наборов сфер заменить на геометрическую дополнительность, т. е. добиться, чтобы сфера S_i^k не пересекала сфер $\tilde{S}_j^k$ с номерами $j \neq i$, а сферу $\tilde{S}_i^k$ пересекала ровно в одной точке, то тогда эти пары дополнительных (друг к другу)

ручек можно было бы устранить. Другими словами, пара таких ручек взаимно уничтожалась бы, и в их окрестности пленка приобрела бы вид прямого произведения. Взаимное уничтожение ручек дополнительных индексов, сепаратрисные сферы которых пересекаются ровно в одной точке, описано в [3] (рис. 17). Для замены алгебраической дополнительности геометрической, т. е. для устранения лишних точек пересечения сепаратрисных сфер, служит прием Уитни.

Шаг второй. Прием Уитни состоит в следующем. Пусть имеются две точки пересечения сфер S и $\tilde{S}$ и пусть индексы пересечения равны $+1$ и -1 (рис. 17(б)). На петлю Уитни, образованную двумя путями, идущими из одной точки пересечения в другую по разным сферам, натягивается вложенный двумерный диск (диск Уитни). Вдоль этого диска одна сфера деформируется так, что стягивается, соскальзывает с другой сферы с помощью изотопии (рис. 17(б)). Для осуществимости этой операции нужно, чтобы петля Уитни была гомотопна нулю в пространстве $X = M \setminus (S \cup \tilde{S})$, т. е. в дополнении к сферам. Это заведомо обеспечивается следующим условием: $\pi_1 X = 0$. Если размерность объемлющего многообразия больше четырех, то это условие выполнено, поскольку $\pi_1(M) = 0$ и коразмерность $S \cup \tilde{S}$ велика. В размерности 4 эти условия не выполнены. Это и есть первая причина, препятствующая выполнению приема Уитни прежними, известными ранее средствами.

Если $\pi_1(X) = 0$, то двумерный диск Уитни с особенностями (сингулярный диск) на указанную петлю натянуть можно. В случае большой коразмерности его можно сделать вложенным, применив малое шевеление диска и используя теоремы общего положения. В размерности 4 этих соображений, вообще говоря, недостаточно. В этом заключается следующая, вторая причина отказа приема Уитни. Впрочем, удобно назвать эту причину четвертой, по соображениям, которые будут ниже описаны. Для доказательства теоремы нужно также добиться того, чтобы диски Уитни, натянутые на разные петли Уитни, не пересекались друг с другом. В этом случае устранение лишних точек пересечения сфер можно осуществить одновременно. В противном случае, т. е. если диски Уитни пересекаются, при первой операции устранения пары точек может нарушиться, например, односвязность дополнения X . И опять, добиться отсутствия пересечений дисков Уитни можно в высоких размерностях, но, вообще говоря, прежние приемы не позволяют этого сделать в размерности 4. Эту причину, препятствующую выполнению приема Уитни, назовем третьей. Наконец, вторая причина заключается в том, что диск Уитни должен «хорошо примыкать» к обеим сферам. В высоких размерностях этого легко добиться, а в размерности 4 возникает некоторое целочисленное препятствие, мешающее хорошей (нулевой) примыкаемости. Суть этого эффекта состоит в следующем. Пусть диск Уитни является уже вложенным в дополнение к сферам.

Тогда его нормальное расслоение (с двумерными слоями), суженное на замкнутую петлю Уитни, имеет две тривиализации. Одна тривиализация индуцируется тем, что нормальное расслоение имеет базой диск, а потому тривиально, является прямым произведением. Вторая тривиализация определяется векторными полями, нормальными к диску и касающимися сферы S на одной половине петли Уитни и сферы $\tilde{S}$ — на другой ее половине (рис. 18). Эти две

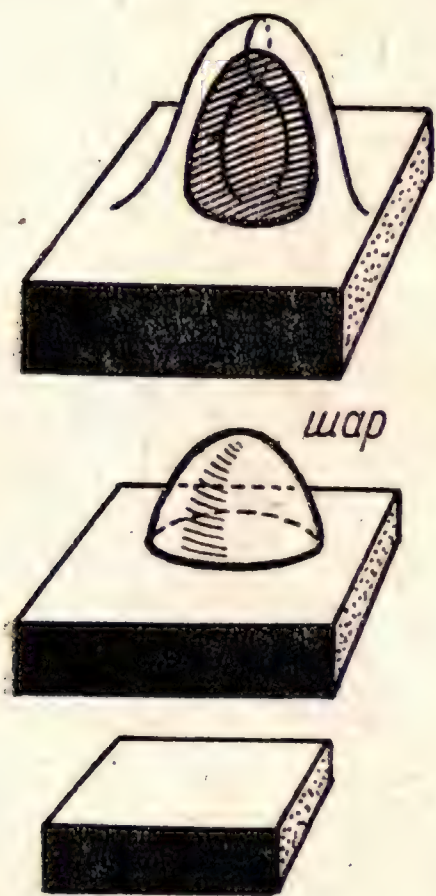


Рис. 17(а)

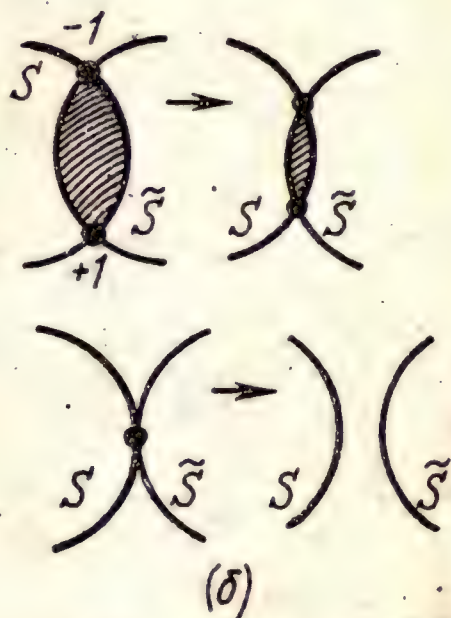


Рис. 17(б)



Рис. 18

тривиализации отличаются на отображение окружности S^1 в группу $SO(2)$, т. е. на элемент из группы $\pi_1(SO(2)) = \mathbb{Z}$. Итак, причины неэффективности приема Уитни (в прежнем его понимании) в размерности 4 таковы:

1) неодносвязность дополнения к сферам; 2) плохая при-
мыкаемость двумерного диска Уитни к сферам; 3) пересечение
разных дисков Уитни, натянутых на разные петли Уитни;
4) самопересечение дисков Уитни.

Оказывается, эти причины можно устранить. Для ликвида-
ции первого препятствия служит прием Кассона. Вот в чем он
заключается. Из сущности задачи видно, что мы имеем право
заменять одно из семейств сфер на изотопные. Оказывается,
для аннулирования фундаментальной группы дополнения (если
она была нетривиальна) достаточно изотопий весьма специаль-
ного вида. Один из участков сферы S вытягивается в виде
«языка», «пальца» вдоль некоторого пути и протыкает вторую
сферу $\tilde{S}$. При этом появляются две новые точки пересечения
сфер. Индексы этих точек равны $+1$ и -1 (рис. 19). Любо-
пытно отметить, что, хотя нашей конечной целью является уст-
ранение лишних точек пересечения, однако первым шагом
в направлении решения этой задачи является добавление но-
вых точек пересечения, т. е. усложнение картины.

Утверждается, что при описанной операции добавления па-
ры точек фундаментальная группа дополнения уменьшается.
Более точно, имеет место равенство: $\tilde{\pi}_1 = \pi_1 / \omega$, где $\tilde{\pi}_1$ — новая

фундаментальная группа, π_1 — прежняя группа, а ω — некоторый элемент группы π_1 . Ясно, что достаточно проанализировать события, происходящие в малой окрестности той точки сферы $\tilde{S}$, около которой «язык» протыкает сферу $\tilde{S}$ (рис. 20). Ясно, что до того момента, когда «язык» протыкает сферу, дополнение в шаре к двум непересекающимся дискам было гомотопически эквивалентно букету двух окружностей, т. е. восьмерке. После того как «язык» проткнул сферу, дополнение к двум пересекающимся дискам стало гомотопически эквивалентно склейке двух торов по координатному кресту. Таким образом, при протыкании сферы происходит факторизация фундаментальной группы дополнения по коммутатору

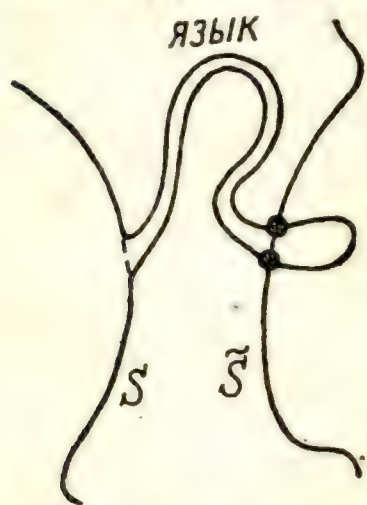


Рис. 19

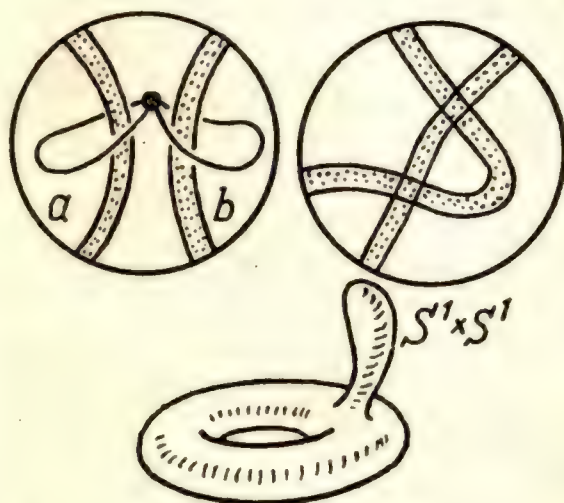


Рис. 20

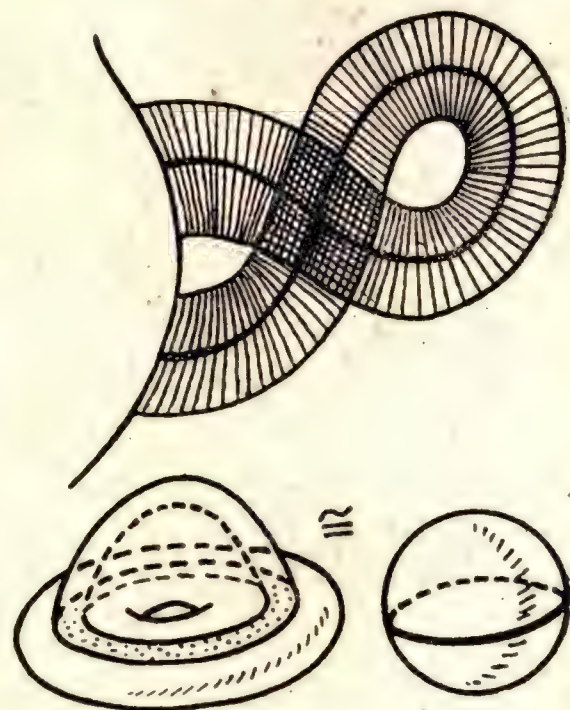


Рис. 21

ω путей a и b (рис. 20). Ясно, что после нескольких таких операций мы полностью уничтожим фундаментальную группу дополнения, так как группа π_1 с самого начала была совершенной. Это следует из того, что $H_1(X) = 0$, что в свою очередь следует из гомотопичности нулю меридианов сфер.

Устранения второго препятствия (см. выше) можно добиться более или менее обычными топологическими средствами. Детали мы здесь опускаем.

Для устранения третьего препятствия на пути применения приема Уитни следует поступать следующим образом. Сначала натягиваем один диск Уитни, потом с помощью приема Кассона выпускаем языки из диска и делаем дополнение к диску односвязным. Следующий диск будем натягивать (на следующую петлю Уитни) уже в дополнение к первому диску, которое односвязно. Очевидно, что этим мы добиваемся отсутствия пересечений у разных дисков Уитни.

После устранения первых трех причин, мешающих реализации приема Уитни, мы получаем несколько не пересекающихся погруженных дисков Уитни с хорошим примыканием. Рассмотрим для простоты случай, когда имеется один диск. Выделим некоторую замкнутую окрестность диска, расслоенную на нормальные двумерные диски малого радиуса. Иногда ее называют витой ручкой. Витая ручка полностью определяется чис-

лом положительных и числом отрицательных двойных точек самопересечения диска. Топологически эта ручка гомеоморфна связной сумме нескольких экземпляров $S^1 \times D^3$. Каждое такое прямое произведение возникает вследствие появления замкнутых петель, начинающихся и кончающихся в двойных точках (рис. 21). На рисунке (размерности два) такая ручка гомеоморфна $S^1 \times I$. Петли, выходящие из двойных точек по одной ветви и возвращающиеся по другой, называются петлями Уитни.

Ясно, что их можно вытеснить на край витой ручки. Отметим важное наблюдение: если каждую петлю Уитни заклеить ручкой индекса 2, то получится шар, т. е. настоящая ручка. Аналог этой операции в трехмерном случае см. на рис. 21. Для того чтобы эту заклейку реализовать внутри многообразия, нужно иметь двумерные диски, натянутые на петли Уитни, вложенные в многообразие и хорошо прилегающие к витой ручке. При этом эти диски должны быть попарно-непересекающимися. Как и выше, можно удовлетворить всем этим условиям, кроме условия вложенности дисков. Заклеиваем петли Уитни погруженными двумерными дисками и берем их замкнутые окрестности. Получаем, следовательно, второй слой витых ручек. Если бы их окружности Уитни можно было заклеить внутри объемлющего многообразия вложенными дисками, удовлетворяющими перечисленным ранее условиям, то эти витые ручки превратились бы в настоящие ручки. В этом случае и исходная витая ручка стала бы настоящей. Однако в общем случае добиться этого мы пока не можем, поэтому вынуждены повторить процесс и построить третий слой витых ручек и т. д. Объединение всех таких витых ручек (построенных в результате описанного процесса) называется ручкой Кассона (рис. 22). Отбросив боковую трехмерную поверхность ручки Кассона, получим открытую ручку Кассона. Фридман доказал, что любая открытая ручка Кассона гомеоморфна прямому произведению $D^2 \times \mathbb{R}^2$, т. е. настоящей ручке.



Рис. 22

Из этой теоремы сразу следует теорема об h -кобордизме в размерности 4. Дело в том, что ось $D^2 \times \{0\}$ ручки $D^2 \times \mathbb{R}^2$ является искомым, правда, только топологическим диском Уитни, вложенным в объемлющее многообразие. Это означает, что прием Уитни работает (но портит гладкость) в размерности 4. Дальше проходит обычное доказательство теоремы об h -кобордизме, как, впрочем, и теоремы об s -кобордизме.

Вкратце опишем доказательство теоремы о гомеоморфности

открытой ручки Кассона обычной ручке. В открытой ручке Кассона $\mathring{K}$ строится открытое подмножество G , имеющее вид, показанный на рис. 23. Каждая пара симметричных отрезков типа AB и $A'B'$ изображает дополнение в S^3 к «толстому» зацеплению Уайтхеда в S^3 , а ветвление происходит по торах (рис. 23(б)). В стандартной ручке $D^2 \times \mathbb{R}^2$ строится открытое подмножество G_1 , показанное на рис. 24. Каждый квадрат

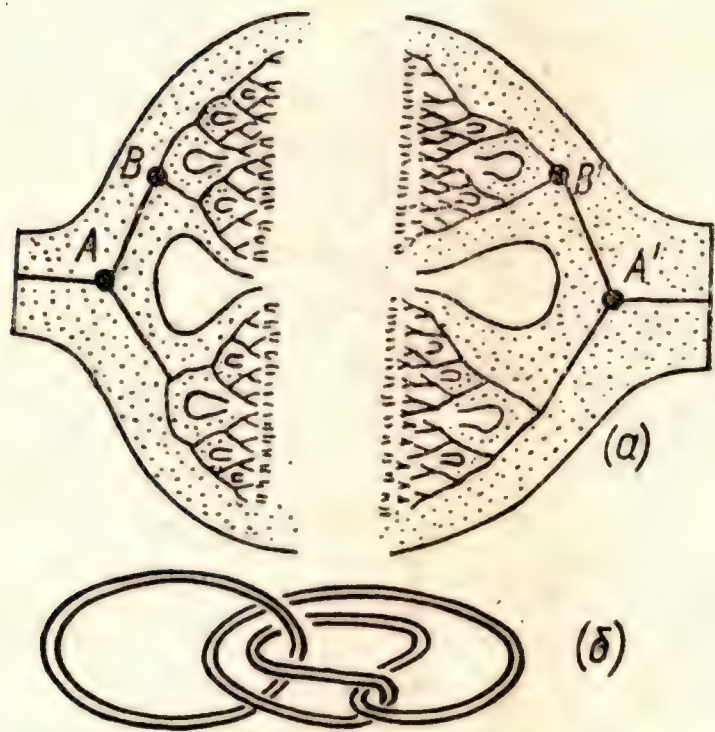


Рис. 23

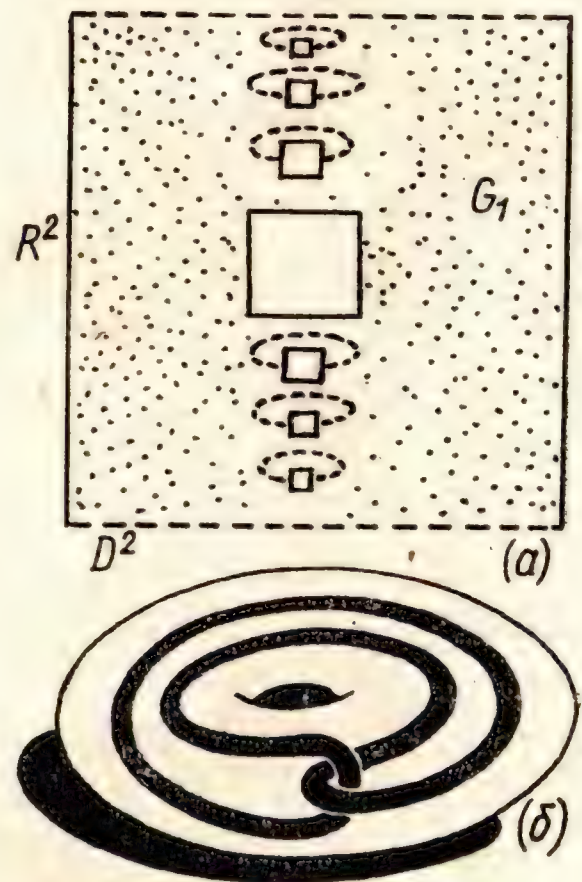


Рис. 24

представляет из себя произведение заполненного тора, лежащего в $D^2 \times \partial D^2 = D^2 \times S^1$, как показано на рис. 24(б), на отрезок. Это прямое произведение «продавлено» внутрь многообразия $D^2 \times \mathbb{R}^2 = D^2 \times \text{Int } D^2$ на нужную глубину. Легко видеть, что множества G и G_1 гомеоморфны. Это — заготовка для построения искомого гомеоморфизма ручки Кассона $\mathring{K}$ с настоящей ручкой $\mathring{P} = D^2 \times \mathbb{R}^2$.

На параллель торам натягиваем вложенный диск (пунктир на рис. 24) и сжимаем каждый тор с диском в точку. Получаем некоторое пространство X и отображение $g: \mathring{P} \rightarrow X$. Отображение $f: \mathring{K} \rightarrow X$ строится по суперпозиции на G и сжиманием компонент дополнения к G в соответствующие точки. При этом следует следить за тем, чтобы конструируемое отображение было непрерывным. Получаем следующую диаграмму: $\mathring{K} \xrightarrow{f} X \xleftarrow{g} \mathring{P}$. Техника, разработанная Бингом, позволяет доказать, что отображение g можно заменить на гомеоморфизм. Техника, разработанная Фридманом, позволяет доказать, что отображение f также можно заменить гомеоморфизмом. Остановимся более подробно на свойствах отображения g . Возьмем

в квадрате отрезок и сожмем его в точку (рис. 25). Очевидно, получим квадрат. Если теперь не в квадрате, а в ручке P сжать в точку не отрезок, а диск с утолщенным краем (т. е. тор с диском, натянутым на параллель) или несколько таких утолщенных дисков, то получим, конечно, снова ручку P . Принцип сжатия Бинга позволяет доказать это утверждение для счетного числа таких утолщенных дисков. Именно такая ситуация реализуется в нашем случае. Тем самым отображение g заменяется на гомеоморфизм.

Остановимся более подробно на свойствах отображения f . Возьмем отображение сферы S^4 на себя, являющееся «почти гомеоморфизмом». Более точно, при этом отображении прообраз каждой точки (за исключением одной) является точкой. Прообразом же этой одной «особой» точки является произвольный компакт Y . Если особую точку A окружить сферой S^3 , то ее прообраз также является трехмерной сферой и по теореме Шенфлиса ограничивает четырехмерную клетку. Поэтому компакт Y можно зажать в сколь угодно тесную клетку. Такие компакты называются клеточными (рис. 25). Ясно, что клеточный компакт можно сжать в точку внутри сферы, заменив тем самым исходное отображение гомеоморфизмом. Это рассуждение проходит для ручки Кассона и для конечного числа особых точек. Техника, разработанная Фридманом, позволяет перенести рассуждение на случай счетного нигде не плотного множества особых точек. Истоки этой техники лежат в работе М. А. Штанько [57]. Именно такая ситуация реализуется в нашем случае, следовательно, отображение f можно заменить гомеоморфизмом. Тем самым завершается доказательство теоремы.

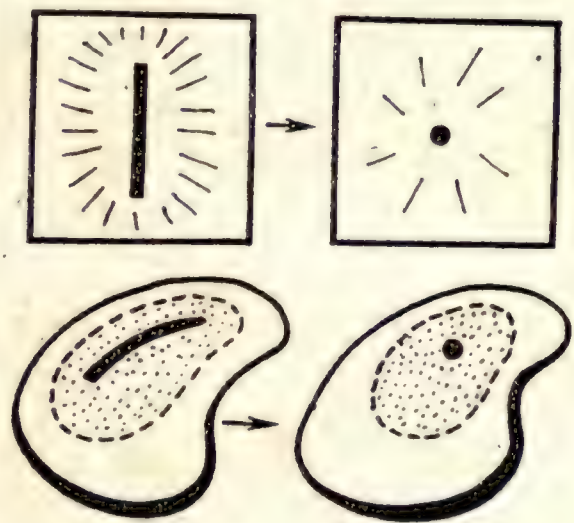


Рис. 25

Гипотеза Пуанкаре тесно связана с классификацией многообразий, поэтому не удивительно, что теорема об h -кобордизме между 4-мерными многообразиями позволила доказать не только гипотезу Пуанкаре в размерности 4, но и классифицировать односвязные 4-мерные многообразия. Вкратце остановимся на этом результате. Уточним класс многообразий, подлежащих классификации. Это замкнутые односвязные топологические почти сглаживаемые 4-мерные многообразия. Почти сглаживаемость означает, что многообразие становится сглаживаемым после удаления из него одной точки. В качестве отношения эквивалентности рассматривается гомеоморфизм. Для классификации нужны два алгебраических инварианта.

Первый инвариант. Это симметричная билинейная целочисленная унимодулярная форма, определяемая индексом пересечения на двумерных гомологиях многообразия, т. е. на груп-

пе $H_2(M^4)$. Другими словами, сопоставляя каждой паре двумерных циклов их индекс пересечения, мы и получаем искомую форму. Следовательно, каждому 4-мерному многообразию M естественным образом отвечает некоторая форма L_M .

Например, многообразию $S^2 \times S^2$ отвечает форма U с матрицей $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; комплексному проективному пространству CP^2 — форма с матрицей (1).

Второй инвариант. Этот инвариант, обозначаемый κ_M , является препятствием к введению векторной структуры на стабильном топологическом касательном расслоении многообразия M . Это препятствие лежит в группе $H^4(M, \mathbb{Z}_2) = \mathbb{Z}_2$. Другими словами, $\kappa_M = 0$, если стабильное касательное расслоение к многообразию M допускает векторную редукцию, и $\kappa_M = 1$ — в противном случае.

Для удобства формулировки теоремы классификации удобно разбить все 4-мерные многообразия на два класса: четные многообразия (с четной формой) и нечетные многообразия (с нечетной формой). Четность формы означает, что на диагонали какой-нибудь (и тогда любой) ее матрицы стоят четные числа, нечетной — если на диагонали имеется хотя бы одно нечетное число.

Теорема классификации 4-мерных многообразий.

1) Сопоставление $M \rightarrow L_M$ индуцирует биекцию между множеством четных многообразий (с указанным выше отношением эквивалентности) и четных форм. 2) Сопоставление $M \rightarrow (L_M; \kappa_M)$ индуцирует биекцию между множеством нечетных многообразий и множеством пар вида: (нечетная форма; число 0, 1).

Отсутствие инварианта κ в четном случае объясняется тем, что здесь он выражается через L_M , т. е. не является независимым параметром. Для доказательства теоремы классификации нужно убедиться в том, что: 1) многообразия с одинаковыми инвариантами гомеоморфны, 2) любые значения инвариантов реализуются.

Отметим, что имеется обобщение теоремы Новикова—Уолла об h -кобордантности компактных гладких 4-мерных многообразий с одинаковыми формами на некомпактный случай. А именно справедливо следующее утверждение: если $L_{M_1} = L_{M_2}$ и, в нечетном случае, $\kappa_{M_1} = \kappa_{M_2}$, то $M_1 \setminus *$ гладко h -кобордантно $M_2 \setminus *$ (в некомпактном смысле с «хорошим поведением на бесконечности»).

По теореме об h -кобордизме в этом случае многообразия M_1 и M_2 гомеоморфны. Тем самым доказано первое утверждение теоремы классификации.

Для того чтобы доказать вторую часть теоремы классификации, нужно доказать реализуемость произвольной формы указанного типа. Для этого потребуются следующие факты:

1) «прием вырезания», 2) существование «колпака Рохлина».

1) Опишем прием вырезания. Пусть задано 4-мерное многообразие M^4 и пусть $M_1 = M \# (S^2 \times S^2)$. Тогда имеет место равенство $L_{M_1} = L_M \oplus U$. Оказывается, верно и обратное, но в топологическом смысле. А именно если $L_{M_1} = L \oplus U$, то существует Тор-многообразие M^4 такое, что $M_1 \stackrel{\text{Тор}}{=} M \# (S^2 \times S^2)$. Это утверждение доказываем так. Реализуем два цикла a, b , задающих форму U , двумя погруженными сферами с одной общей точкой. Для этого сначала реализуем один цикл, потом с помощью описанного выше приема Кассона сделаем его дополнение односвязным и затем уже реализуем в этом дополнении второй цикл, пересекающийся с первым циклом лишь в одной точке (рис. 26). Потом, как и при доказательстве теоремы об h -кобордизме, заклеиваем петли Уитни с помощью описанного бесконечного процесса, т. е. заклеиваем петли витыми ручками. В результате получаем пару ручек Кассона, приклеенных к шару, т. е. пару топологически стандартных ручек, приклеенных к шару, т. е. $(S^2 \times S^2) \setminus D^4$, где D^4 — малый диск. Удалив из многообразия M_1 многообразие $Q = ((S^2 \times S^2) \setminus \text{малый открытый диск})$ и заклеивая край многообразия шаром, мы и получаем искомое многообразие M (рис. 26). Построенное многообразие удовлетворяет нашим требованиям.

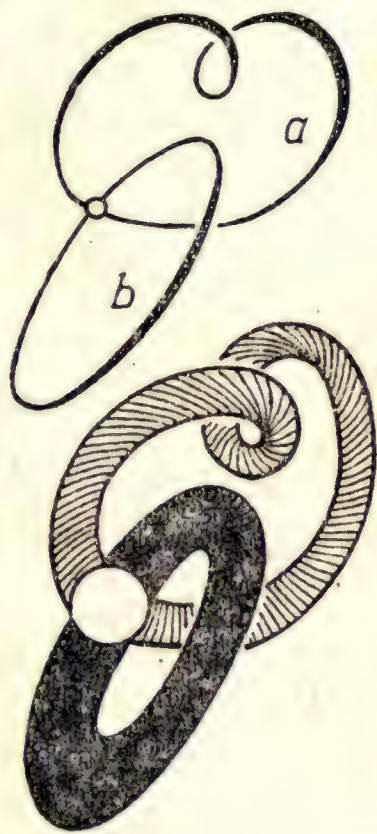


Рис. 26

2) Опишем колпак Рохлина. Колпаком Рохлина называется стягиваемое почти сглаживаемое 4-мерное многообразие M_0^4 , край которого является пространством додекаэдра M_0^3 . Его существование доказывается следующим образом. Возьмем прямое произведение $M_0^3 \times I$. Напомним, что $\pi_1(M_0^3) = \langle x, y \mid (xy)^2 = x^3 = y^5 \rangle$. Сделаем сферическую перестройку по элементу (окружности-циклу) x , удалив его окрестность, гомеоморфную $S^1 \times D^3$, и вклеив вместо нее прямое произведение $D^2 \times S^2$. В результате получим 4-мерное многообразие V . Легко видеть, что $H_2(V, \mathbb{Z}) = \mathbb{Z} \oplus \mathbb{Z}$ с формой U . Одна группа $\mathbb{Z}$ возникает из двумерной сферы, зацепляющей петлю x , вторая группа $\mathbb{Z}$ возникает из приклеенного диска $D^2 \times (*)$ и из пленки, которая осуществляет гомологию петли x к нулю внутри многообразия M_0^3 . Вырезая этот носитель формы U , получаем многообразие W с краем $M_0^3 \cup M_0^3$, имеющее гомотопический тип $S^3 \times I$. Склеим из этих многообразий W «бесконечную трубу» W_∞ , гомотопически эквивалентную $S^3 \times \mathbb{R}^1$ (рис. 27). Допустим, что эту трубу можно рассечь сферой S^3 с воротником (т. е. с окрестностью, являющейся прямым произведением сферы на отрезок) на две части, см. пунктир на рис. 27. Тогда искомый колпак строится так. Нужно взять часть трубы W_∞ , заключенную между сферой (пунктир на рис. 27) и достаточно

далеким, не пересекающим S^3 экземпляром многообразия M_0^3 , и заклеить сферу S^3 шаром (рис. 27). Доказательство существования сферы S^3 , рассекающей трубу, проводится следующим образом. Сначала из трубы W_∞ удаляется прямая l , потом выбирается гладкий диск D , протыкающий ее в одной точке (рис. 27). Тогда $W_\infty \setminus l$ собственно гомотопически эквивалентно $\mathbf{R}^4$, так как $W_\infty \stackrel{p}{\sim} S^3 \times S^1$. По теореме об h -кобордизме для некомпактных многообразий (см. выше) $W_\infty \setminus l$ гладко h -кобордантно $\mathbf{R}^4$ и, следовательно, $W_\infty \setminus l \stackrel{\text{Top}}{\cong} \mathbf{R}^4$. Изобразим пространство $\mathbf{R}^4$ в виде $S^4 \setminus *$. Тогда после компактификации одной точкой мы получим снова сферу S^4 , а из диска без точки $D \setminus (D \cap$

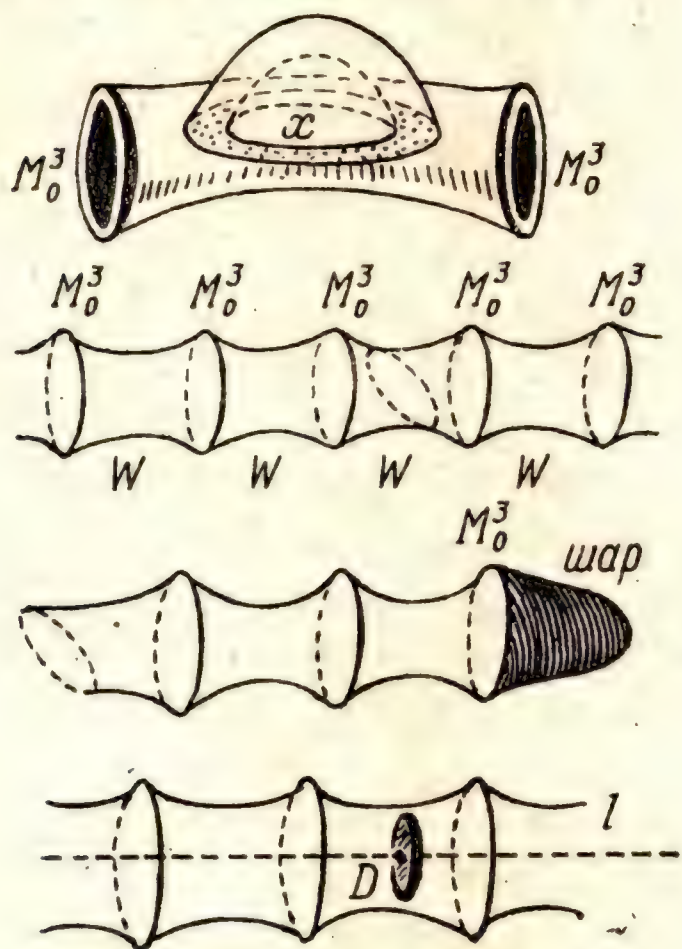


Рис. 27

2	1						
1	2	1					
	1	2	1				
		1	2	1			
			1	2	1		1
				1	2	1	
					1	2	
				1			2

Рис. 28

$\cap l$) получится диск $\bar{D}$ в сфере S^4 , который является локально-плоским везде, кроме, быть может, одной добавленной точки. Здесь можно применить теорему Чернавского—Керби о том, что 3-мерное многообразие в сфере S^4 не может иметь изолированных нелокально-плоских точек. Поэтому диск $\bar{D}$ является локально-плоским, т. е. стандартным. Тогда его можно естественным образом дополнить до сферы. Это означает, что диск D в трубе W_∞ также можно дополнить до искомой сферы, что и требовалось доказать.

Переходим к доказательству теоремы реализации формы. Сначала реализуем форму E_8 с матрицей, показанной на рис. 28. Она замечательна тем, что ее сигнатура равна 8, поэтому ее заведомо нельзя реализовать гладким многообразием. В противном случае это противоречило бы теореме Рохлина, согласно которой сигнатура четного гладкого многообразия обязана делиться на 16.

Возьмем восемь экземпляров пространств касательных рас-
слоений к двумерной сфере S^2 на двумерные диски и склеим
их с помощью так называемого метода плюмбинга по схеме,
показанной на рис. 29. На крае получится трехмерное прост-
ранство додекаэдра M_0^3 [55]. Форма I этого многообразия
задается матрицей E_8 , что следует из его построения. Остается
заклеить многообразие M_0^3 колпаком Рохлина.

Дальнейший процесс реализации затрудняется тем, что
отсутствует классификация всех форм. Однако имеется клас-
сификация неопределенных форм. А именно известно, что лю-
бая четная неопределенная форма имеет вид $kE_8 \oplus mU$, а любая
нечетная неопределенная форма имеет вид $k(+1) \oplus m'(-1)$.

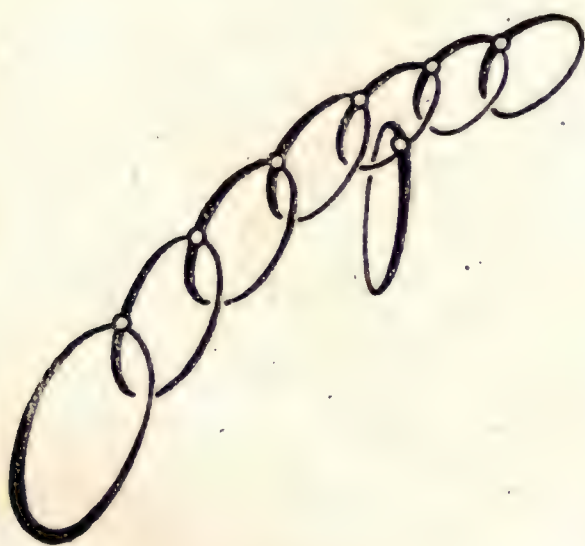


Рис. 29

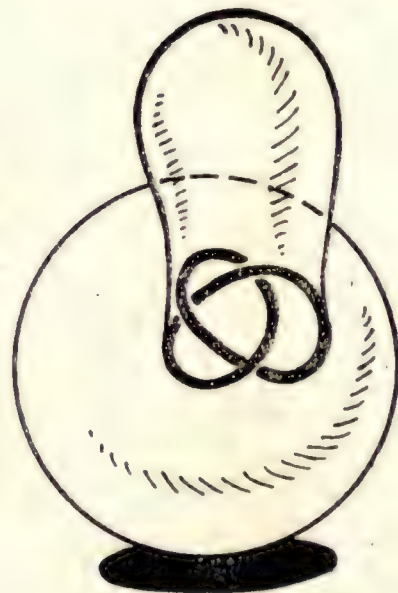


Рис. 30

Эти формы реализовать просто. Для этого достаточно взять
связную сумму многообразий Рохлина с многообразиями $S^2 \times$
 $\times S^2$ в первом случае и взять связную сумму многообразий
 $\pm \mathbb{C}P^2$ во втором случае. Перейдем теперь к реализации опре-
деленных форм.

Пусть есть некоторая форма L . Добавив форму U , получа-
ем неопределенную форму $L \oplus U$. Эту форму реализуем по опи-
санному выше рецепту для неопределенных форм, а затем
с помощью соответствующего вырезания (см. выше) получаем
реализацию формы L .

Во всех рассмотренных случаях реализации нечетных форм
мы получали $\kappa=0$, так как связная сумма проективных прост-
ранств, т. е. $\# \pm \mathbb{C}P^2$, является гладким многообразием, а при
вырезании носителя формы U инвариант κ не меняется.

Для реализации значения инварианта $\kappa=1$ в нечетном слу-
чае нужно построить так называемое «фальшивое проективное
пространство» $\pm \tilde{\mathbb{C}P}^2$. Это многообразие является 4-мерным,
имеет форму (± 1) и значение инварианта $\kappa=1$. Построение
осуществляется так. Возьмем 4-мерный шар, на его крае нари-
суем одномерный узел — трилистник. По этому узлу по осна-
щению $+1$ приклеивается ручка индекса 2 (рис. 30). В ре-
зультате этой операции на крае получившегося многообразия
оказывается пространство додекаэдра [55], т. е. M_0^3 . Его нуж-

но заклеить колпаком Рохлина. Настоящее проективное пространство CP^2 получается аналогичным приемом, но только вместо трилистника нужно взять стандартно вложенную окружность, а на крае получается сфера S^3 . Теперь в разложении $\#CP^2$ достаточно заменить одно из этих многообразий (или нечетное их число) фальшивым проективным пространством. В результате мы получим, что $\kappa=1$. Перейдем теперь к задаче реализации форм гладкими многообразиями. Выпишем все четыре возможности, описывающие типы форм. Зададим их в виде следующей матрицы:

	определенные	неопределенные
четные	I	II
нечетные	III	IV

Относительно типов I и III имеется теорема Дональдсона: из определенных форм гладким многообразием реализуется только форма с единичной матрицей. Относительно форм типа II известно следующее. Известна реализуемость поверхностью Куммера формы $E_8 \oplus E_8 \oplus U \oplus U \oplus U$ и, следовательно, любой формы с числом форм U , не меньшим, чем $3/5$ от числа форм E_8 , причем число форм E_8 должно быть четно. В противном случае сигнатура не делится на 16. Формы типа IV (см. таблицу выше) реализуются все. Итак, получаем следующую матрицу реализуемости форм разных типов:

	определенные	неопределенные
четные	нет	часть
нечетные	одна серия	все

Доказательство теоремы Дональдсона использует современные факты из теории дифференциальных уравнений в частных производных, в частности понятие самодвойственной связности Янга—Миллса, инстантоны.

Оказывается, далее, что из теоремы Дональдсона следует также существование нестандартной гладкости на $\mathbf{R}^4$. Дело в том, что если бы любая гладкая структура на $\mathbf{R}^4$ была стандартной, то алгебраическое расщепление $L_{M_1} = L \oplus U$ можно было бы продублировать геометрическим расщеплением $M_1 = M \# (S^2 \times S^2)$ не только в топологическом, но и в гладком смысле. Это противоречит теореме Дональдсона, поскольку

тогда от поверхности Куммера с формой $E_8 \oplus E_8 \oplus U \oplus U \oplus U$ можно было бы отщепить три формы U и получить гладкое многообразие с формой $E_8 \oplus E_8$. Опишем эту процедуру более подробно.

Итак, предположим, что любая гладкая структура на $\mathbf{R}^4$ является стандартной. Докажем гладкую отщепляемость формы E_8 . Напомним, что отщепление происходит так. В многообразии M_1 отщепляемые циклы реализуются погруженными сферами, пересекающимися лишь в одной точке. Затем производится убивание (заклейка) петель Уитни с помощью приклейки витых ручек. В результате получается открытое множество H в многообразии M_1 , гомеоморфное проколотому многообразию $S^2 \times S^2$. Это построение можно продублировать (в гладком смысле) в стандартном многообразии $S^2 \times S^2$ следующим образом. Нужно взять в многообразии $S^2 \times S^2$ стандартную пару сфер, локально ввести (путем добавления перекрещенной двумерной сферы с одной точкой самопересечения в четырехмерном пространстве) столько положительных и отрицательных точек самопересечения, сколько их имеется на начальном этапе в многообразии M_1 . Затем нужно рассмотреть все двумерные диски, заклеивающие петли Уитни и являющиеся вложенными в многообразие $S^2 \times S^2$, и «испортить» эти диски путем добавления к ним маленьких перекрещенных двумерных сфер (с одной точкой самопересечения) так, чтобы число положительных и отрицательных точек самопересечения совпадало с числом соответствующих точек в соответствующих дисках в многообразии M_1 . В результате получится открытое множество H_1 в многообразии $S^2 \times S^2$, диффеоморфное (!) множеству H . Если теперь из $S^2 \times S^2$ удалить сжатое (до множества, гомеоморфного $(S^2 \times S^2) \setminus (\text{замкнутый шар})$) множество H_1 , то получится открытое подмножество, гомеоморфное $\mathbf{R}^4$. Так как гладкость, наследуемая из многообразия $S^2 \times S^2$, является на этом $\mathbf{R}^4$ стандартной, то ее можно продолжить до гладкой структуры на S^4 . Это означает, что гладкость на множестве $M_1 \setminus (\text{сжатое } H)$ можно продолжить на заклеивающий диск и получить гладкое многообразие M такое, что $M_1 = M \# U$ в гладком смысле (рис. 31). Тем самым доказательство завершается.

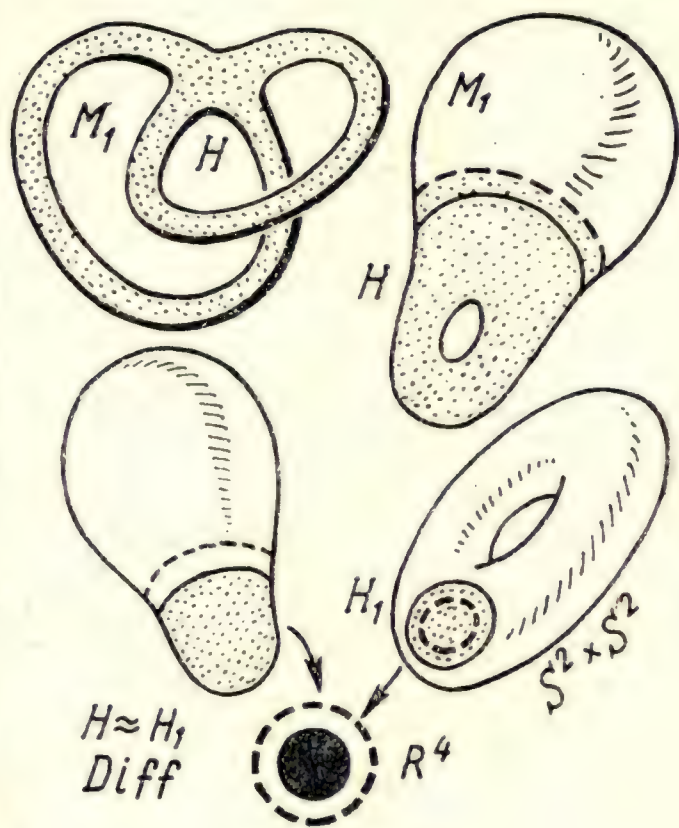


Рис. 31

Аналогичным образом можно доказать существование не-стандартной гладкой структуры на любом гладком 4-мерном многообразии M^4 , из которого выколота одна точка. Для этого

нужно приклеивать $S^2 \times S^2$ гладким образом, а отрезать его — не гладко.

Итак, подведем некоторые итоги.

1) Любое метризуемое топологическое четырехмерное многообразие, гомотопически эквивалентное стандартной четырехмерной сфере, гомеоморфно этой сфере. 2) Любое топологическое четырехмерное многообразие, собственно гомотопически эквивалентное $\mathbb{R}^4$, гомеоморфно $\mathbb{R}^4$. 3) Любое трехмерное топологическое многообразие, целочисленные гомологии которого изоморфны целочисленным гомологиям стандартной трехмерной сферы, является краем стягиваемого топологического четырехмерного многообразия. 4) Замкнутые односвязные топологические почти сглаживаемые многообразия (т. е. допускающие гладкую структуру на дополнении к точке) допускают классификацию. 5) Существует гладкое четырехмерное многообразие, гомеоморфное, но не диффеоморфное стандартному четырехмерному евклидову пространству.

В теореме классификации важную роль играет форма L_M многообразия M . Если M — замкнутое односвязное четырехмерное многообразие, то его целочисленные гомологии и когомологии имеют следующий вид:

$$\begin{array}{cccccc} & i=0 & 1 & 2 & 3 & 4 \\ H_i(M) = \mathbb{Z} & 0 & \mathbb{Z}^q & 0 & \mathbb{Z} \\ H^i(M) = \mathbb{Z} & 0 & \mathbb{Z}^q & 0 & \mathbb{Z} \end{array}$$

Здесь q — второе число Бетти многообразия M . Если a и b — элементы группы $H_2(M, \mathbb{Z})$, то определен их гомологический индекс пересечения $a \cdot b = b \cdot a \in \mathbb{Z}$. В силу двойственности Пуанкаре $H_2(M) \approx H^2(M)$. Таким образом, возникает симметричная билинейная невырожденная форма $L_M(a, b) = a \cdot b$; $L_M: H_2(M) \times H_2(M) \rightarrow \mathbb{Z}$, которую можно задать симметрической целочисленной унимодулярной матрицей. Эта форма полностью определяет гомологическую структуру замкнутого односвязного четырехмерного многообразия.



§ 1. Простейшие свойства минимальных поверхностей

1. Физические опыты Плато и способы получения мыльных пленок.

Прежде чем приступать к математическому изучению минимальных поверхностей, дадим краткое описание тех их основных свойств, которые можно продемонстрировать на языке наглядной геометрии. Когда бельгийский физик, профессор физики и анатомии Жозеф Плато (Joseph Plateau, 1801—1883) начал опыты по изучению конфигурации мыльных пленок, он вряд ли предполагал, что они послужат толчком к возникновению значительного научного направления, бурно развивающегося до настоящего времени и известного сегодня под общим названием «проблема Плато». Некоторые опыты, которые когда-то проводил Плато, чрезвычайно просты и хорошо знакомы читателю; так, вряд ли кто-нибудь не развлекался в свое время выдуванием мыльных пузырей через трубку или конструированием разнообразных мыльных пленок, затягивающих проволоочный контур. Хорошо известно, что если в мыльную воду опустить замкнутый проволоочный контур (например, гомеоморфный окружности) и затем осторожно извлечь его, то на контуре повиснет красивая радужная мыльная пленка, ограниченная контуром. Размеры пленки могут быть довольно значительны, но чем больше пленка, тем легче и быстрее она лопается под воздействием силы тяжести. Если, напротив, размеры контура сравнительно невелики, то можно пренебрегать силой тяжести при изучении некоторых важных свойств мыльной пленки. Этим обстоятельством мы и будем в дальнейшем постоянно пользоваться. Минимальные поверхности являются математическим объектом, достаточно хорошо моделирующим физические мыльные пленки. Обратно, многие глубокие свойства минимальных поверхностей проявляются чрезвычайно наглядно в простых физических опытах с мыльными пленками. Существует много математических направлений, возникших из конкретных физических и прикладных задач, однако далеко не все они столь тесно, как задача Плато, связаны с таким количеством разнообразных по своему аппарату математических теорий. В рамках теории минимальных поверхностей переплетаются теории дифференциальных уравнений, групп Ли, гомологий и когомологий, бордизмов и т. д.

В настоящем параграфе мы проиллюстрируем на простых примерах некоторые основные понятия, технику и результаты, разработанные в различные периоды развития задачи Плато. Поскольку точные математические конструкции в этой области требуют иногда весьма разветвленной, рафинированной и нетривиальной техники, то мы пока ограничимся лишь наглядными построениями, опуская громоздкие вычисления. В частности, мы в значительной мере опустим обобщения на многомерный случай (так называемую многомерную проблему Плато). Некоторые данные мы возьмем из [11] и [19].

Значительное внимание уделяли математическому изучению свойств мыльных пленок еще в XVIII в. Эйлер и Лагранж (задача нахождения поверхности наименьшей площади с за-

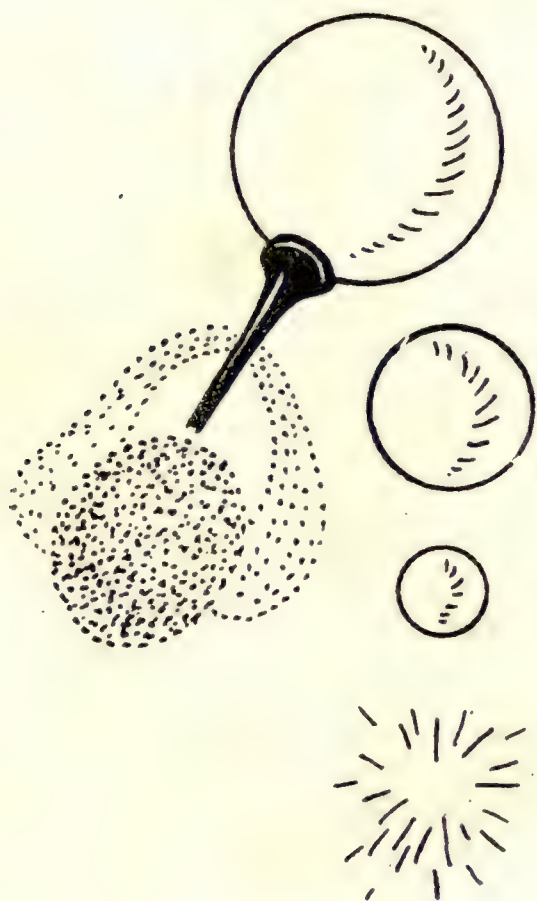


Рис. 1

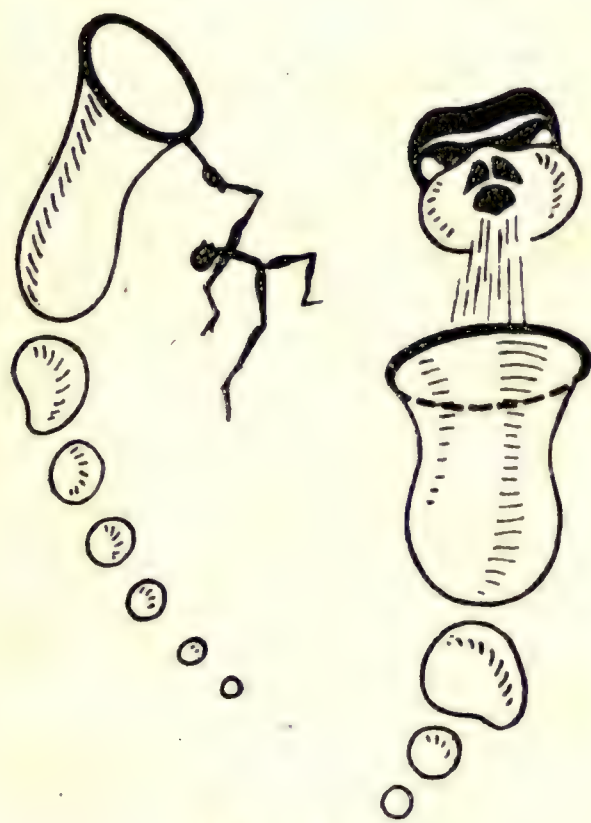


Рис. 2

данной границей). Точные решения для некоторых весьма специальных граничных контуров были найдены Риманом, Шварцем, Вейерштрассом уже в XIX в. Теория минимальных поверхностей возникла из рассмотрения двух сортов мыльных пленок: мыльных пузырей и мыльных пленок, затягивающих некоторый проволоочный контур.

Первый стандартный способ получения мыльных пузырей заключается в их выдувании через трубку, предварительно опущенную в мыльную воду (рис. 1). Для стабилизации пузырей (которые обычно быстро лопаются) в воду добавляют глицерин. Такие мыльные пленки-пузыри характеризуются тем, что они образованы и удерживаются внутренним давлением газа (воздуха), заключенного внутри пленки. Шаровая форма пузыря легко объясняется тем, что эта форма обеспечивает наименьшую площадь минимальной поверхности при фиксиро-

ванном объеме, ограниченном этой пленкой. Вторым способ получения мыльных пузырей изображен на рис. 2. Для этого нужно взять замкнутый контур (например, окружность), опустить его в мыльную воду, извлечь обратно и быстро переместить контур в направлении, ортогональном плоскости контура (для простоты считаем, что контур является плоской кривой). Тогда пленка прогибается под давлением набегающего потока воздуха и от нее отрываются один за другим несколько мыльных пузырей. Ясно, что того же эффекта можно добиться, направляя струю воздуха на мыльную пленку, висящую на неподвижном контуре.

Другой сорт мыльных пленок получается, когда мы извлекаем из мыльной воды проволочный контур и фиксируем его в пространстве, избегая каких-либо резких перемещений. Тогда на контуре возникает устойчивая мыльная пленка, обычно не имеющая в своей структуре мыльных пузырей описанного

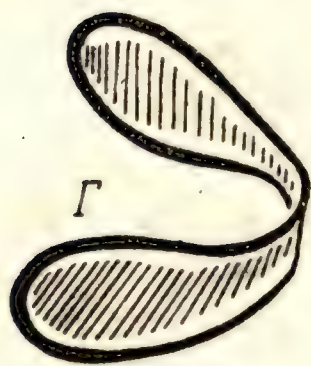


Рис. 3

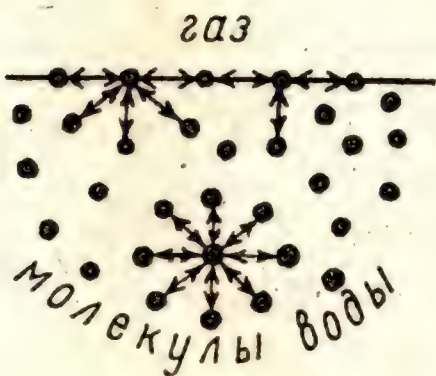


Рис. 4



Рис. 5

типа (возникающих при наличии разности давлений газа внутри и снаружи пленки). Эта пленка имеет в качестве границы исходный проволочный контур (рис. 3). Сразу же отметим, что этот второй тип мыльных пленок (ограниченных контуром) является пока основным объектом, подвергающимся изучению математическими методами. Связано это с тем, что в большом количестве прикладных задач минимальная поверхность прикреплена в какой-то фиксированной границе, как это имеет место, например, в теории мембран.

2. Физические принципы, лежащие в основе образования мыльных пленок.

Физический принцип, формирующий мыльные пленки, регулирующий их поведение, локальные и глобальные свойства, чрезвычайно прост: физическая система сохраняет определенную конфигурацию только в том случае, когда она не может легко изменить ее, заняв положение с меньшим значением энергии. В нашем случае энергия поверхности (мыльной пленки), часто описываемая в терминах поверхностного натяжения жидкости, обусловлена наличием сил притяжения между отдельными молекулами и несбалансированностью этих сил на границе поверхности. Наличие таких несбалансированных сил влечет за собой интересный эффект: жидкая пленка превра-

щается в эластичную поверхность, стремящуюся минимизировать свою площадь и, следовательно, минимизировать энергию натяжения, приходящуюся на единицу площади. При этом мы, как и раньше, пренебрегаем силой тяжести (в случае пузырей и в случае пленок с границей) и давлением воздуха (в случае пленок с границей).

Рассмотрим несколько подробнее изменение свойств поверхностной пленки при добавлении в воду мыла. На рис. 4 условно изображена граница раздела двух сред: воды и воздуха. Черными кружками условно изображены молекулы воды.

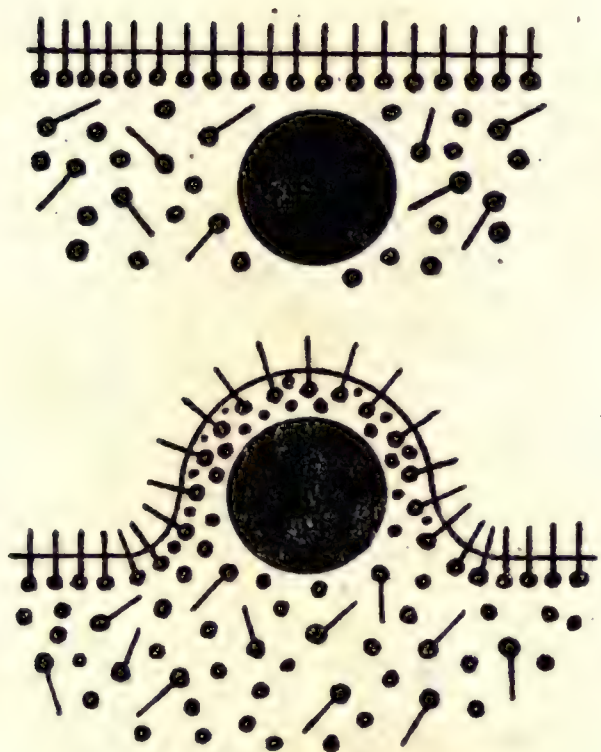


Рис. 6(а)

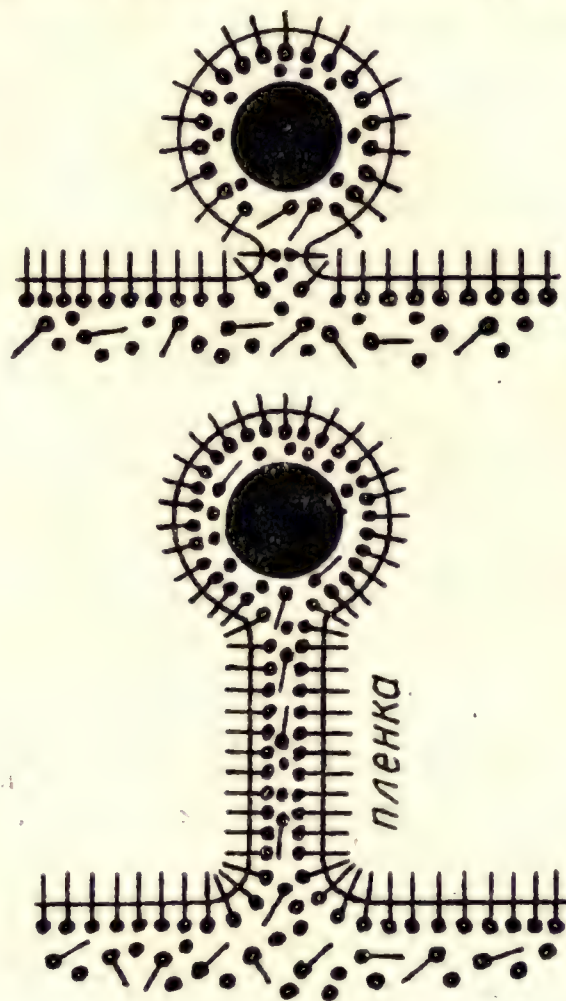


Рис. 6(б)

Некоторые из них находятся в воздухе, поскольку жидкость испаряется, но этим эффектом мы пренебрегаем. Стрелками с двумя концами обозначены силы взаимного притяжения, действующие между полярными молекулами воды, характеризующимися, как известно, асимметричным расположением электрического заряда. Отчетливо видно, что именно эти силы взаимного притяжения вызывают появление поверхностного натяжения на границе сред. Именно таким образом и формируются свойства поверхностной пленки жидкости. В отличие от молекул воды молекулы мыла образованы длинными тонкими неполярными углеводородными цепями с полярной кислородной группой на одном из концов цепи. Когда эти молекулы добавлены в воду, они устремляются к ее поверхности и заполняют поверхность жидкости равномерным слоем (рис. 5). При этом каждая молекула мыла ориентирована на поверхности своим неполярным концом наружу. Оттесняя молекулы воды внутрь жидкости, молекулы мыла уменьшают тем самым поверхностное натяжение на границе раздела двух сред. Ока-

зывается, это обстоятельство придает поверхностной пленке дополнительную эластичность, проявляющуюся особенно в тот момент, когда мы опускаем, а затем извлекаем из мыльной воды тонкий проволочный контур. Этот процесс показан на рис. 6. В самом деле, когда проволока, поднимаясь из жидкости, достигает ее поверхности (рис. 6), то поверхность вспучивается, обволакивая контур (на рисунке изображено его сечение). Это происходит вследствие того, что около проволоки число молекул мыла на поверхности временно уменьшается (рис. 6), т. е. падает их число на единицу площади пленки, а следовательно, в этом месте увеличивается поверхностное натяжение. Оно создается прорывающимися снизу на освободившиеся места молекулами воды. Это придает поверхностной пленке в непосредственной близости от проволоки дополнительную гибкость, эластичность, приводящую в итоге к появлению мыльной пленки после окончательного извлечения контура из воды. Процесс образования пленки, висящей на контуре, показан на рис. 6. По мере подъема контура вверх гибкая поверхностная пленка не прорывается проволокой, а обволакивает ее и тянется вслед за ней, вытягиваясь из жидкости. В результате получается мыльная пленка, затягивающая контур. Сила тяжести ограничивает размер минимальной поверхности, и при значительном удалении некоторых частей пленки от граничного контура пленка лопаётся: сил поверхностного натяжения становится недостаточно для удерживания пленки в равновесии.

3. Экстремальные свойства мыльных пленок и минимальность их площади. Свойства границы раздела двух сред.

Именно в силу описанного механизма математической моделью мыльной пленки может служить так называемая минимальная поверхность, т. е. поверхность, имеющая наименьшую возможную площадь среди всех поверхностей, имеющих ту же границу. Ясно, что форма поверхности и ее свойства в значительной мере определяются конфигурацией граничного контура — края поверхности. По-видимому, само понятие наименьшей площади было впервые введено в геометрию Архимедом, который не только отметил, что прямая линия реализует кратчайшее расстояние между двумя точками, но и хорошо понимал, что плоскость также может быть охарактеризована в терминах минимальной площади. Эти представления, естественно, возникли из повседневной практики, когда, например, для изготовления барабана требовалось сильно натянуть кожу на деревянный остов. Это позволило связать представление о поверхностном натяжении со свойством поверхности иметь минимальную площадь. Вообще, история развития правильных представлений о поверхностном натяжении чрезвычайно интересна. Здесь мы только коснемся некоторых этапов этого пути. По-видимому следующий наиболее существенный прогресс в развитии этих представлений (после длительного перерыва)

был достигнут Бойлем, который в 1676 г. заинтересовался формой капель жидкости. Обратив внимание на то обстоятельство, что капли дождя имеют приблизительно круглую форму, Бойль решил найти зависимость между размером капли и ее формой. При этом следовало поставить довольно тонкий по тем временам эксперимент, чтобы поведение капель можно было изучать на протяжении достаточно длительного времени без их разрушения. Для этих целей Бойль наливал в стеклянный сосуд две жидкости: раствор K_2CO_3 (сравнительно тяже-

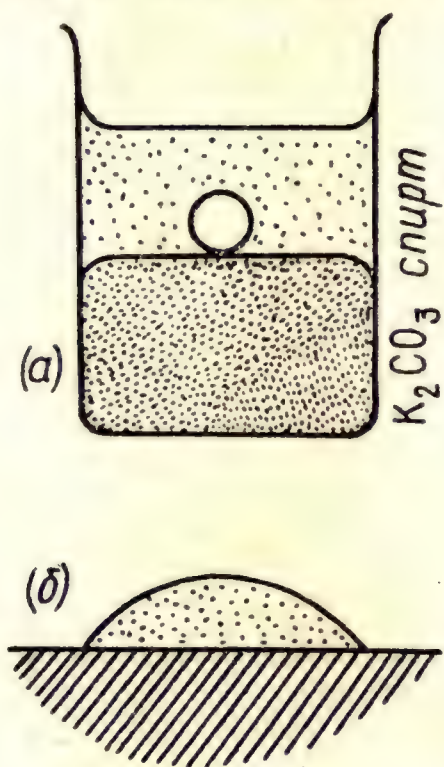


Рис. 7

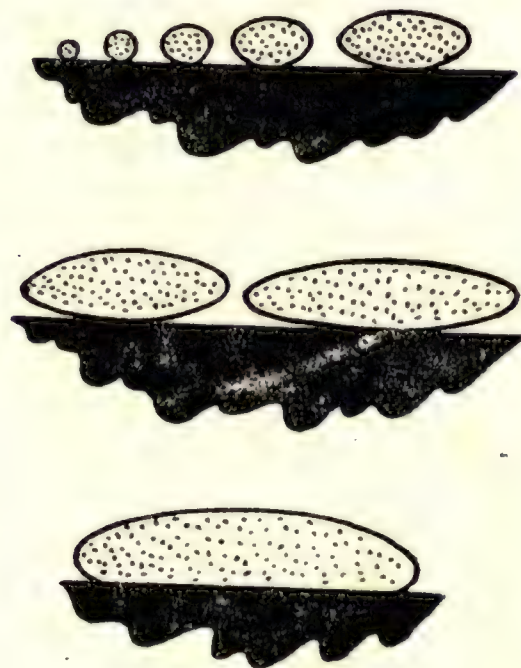


Рис. 8

лая, плотная жидкость — концентрированный раствор карбоната калия, поташа) и спирт (легкая жидкость). Когда жидкости отстоялись, между ними образовалась хорошо различимая поверхность раздела двух сред (рис. 7). Затем на эту поверхность раздела осторожно помещалась капля третьей жидкости промежуточной плотности и несмешиваемой с обеими жидкостями, уже присутствующими в сосуде. Эта капля находилась в равновесии, будучи погружена в спирт и касаясь в своей нижней точке поверхности K_2CO_3 . Эта третья жидкость (масло) подбиралась таким образом, чтобы она не смачивала поверхность раздела сред, т. е. чтобы капля не растекалась по поверхности раздела и не имела бы формы, показанной на рис. 7(б). Затем на поверхность раздела помещались капли разного размера, начиная с небольших капель, после чего их размер постоянно увеличивали. Оказалось, что с ростом размера капли ее форма начинает меняться (рис. 8). Визуально сплющивание капель становилось заметным для капель размером с горошину. Конечно, сегодня нам известно, что сплющивание капель происходит в любом случае, даже для капель малого размера при наличии силы тяжести. Однако для того, чтобы это заметить, нужно было обладать достаточно точными приборами, которыми Бойль в 1676 г., конечно, не располагал. Кроме того, следует помнить, что современные представления

о том, что форма капли должна соответствовать наименьшей энергии, еще только начинали развиваться в XVII в. Большинство процессов такого рода объяснялось тогда в терминах флогистона, что, конечно, еще более затрудняло понимание истинных механизмов, определяющих форму капель. И наконец, в действительности форма капли минимизирует комбинацию поверхностной энергии и силы тяжести, что при малых размерах капли существенно усложняет математическую сторону вопроса. В 1751 г. Сегне (Segner) отчетливо осознал это физическое обстоятельство и понял, что для аккуратного изучения поверхностного натяжения следует попытаться устранить, насколько это возможно, влияние силы тяжести. Он частично добился этого, изучая свободно падающие капли, а также капли, погруженные в жидкость той же плотности. При этом заранее обеспечивалась несмешиваемость жидкости капли и объемлющей жидкости. В частности, ему удалось показать, что сфера имеет наименьшую площадь среди всех замкнутых поверхностей, ограничивающих фиксированный объем. Сегне, по-видимому, впервые понял истинную роль поверхностного натяжения во всех этих процессах. Однако дальнейшее развитие этих идей (включая идею поверхностной энергии) существенно тормозилось доминированием идеи флогистона, не желавшей сдавать позиций.

4. Граница раздела двух сред, находящихся в равновесии, является поверхностью постоянной средней кривизны.

Важный шаг в понимании внутренней геометрии поверхностей раздела сред был сделан Пуассоном в 1828 г., когда он показал, что поверхность раздела двух сред, находящаяся в равновесии (при условии, что мы пренебрегаем силой тяжести), является поверхностью постоянной средней кривизны. Остановимся более подробно на этом важном результате, изложив его в современной терминологии с использованием понятий современной дифференциальной геометрии. Здесь для читателя может оказаться полезным предварительное знакомство с некоторыми основными свойствами второй квадратичной формы двумерной поверхности, вложенной в трехмерное евклидово пространство. Соответствующие факты см., например, в [1, 5].

Пусть M^2 — двумерная гладкая поверхность в $\mathbb{R}^3$, где $\mathbb{R}^3$ отнесено к декартовым координатам x, y, z . Пусть поверхность задается радиус-вектором $r = r(u, v) = (x(u, v), y(u, v), z(u, v))$, где параметры u и v изменяются в некоторой области на евклидовой плоскости, причем определяют регулярные координаты в окрестности точки P на поверхности. Пусть $n(u, v)$ — единичный вектор нормали к поверхности в точке $P = P(u, v)$. Рассмотрим квадратную симметрическую матрицу Q , составленную из чисел $q_{ij} = \langle r_{u_i, u_j}, n \rangle$, где r_{u_i, u_j} — вторые частные производные радиус-вектора по переменным u и v (обозначенным здесь как u_1 и u_2 соответственно). Через $\langle, \rangle$ обозначено

евклидово скалярное произведение в $\mathbb{R}^3$. Если a и b — касательные векторы к поверхности в точке P , то выражение $Q(a, b) = \sum_{i,j} q_{ij} a_i b_j$, где $a = (a_1, a_2)$, $b = (b_1, b_2)$, является би-

линейной формой, определенной на касательной плоскости к поверхности и называемой второй квадратичной формой поверхности. Ясно, что вторая форма зависит от способа вложения поверхности и при ее деформации форма также будет, вообще говоря, изменяться. В двумерном случае, который нас здесь и интересует, вторая квадратичная форма записывается обычно в виде $Ldu^2 + 2Mdudv + Ndv^2$, где $L = L(u, v) = \langle r_{uu}, n \rangle$, $M = M(u, v) = \langle r_{uv}, n \rangle$, $N = N(u, v) = \langle r_{vv}, n \rangle$. Здесь L, M, N — гладкие функции от координат на поверхности. Наряду со второй квадратичной формой рассмотрим также первую квадратичную форму на поверхности, определяющую риманову метрику, индуцированную на поверхности объемлющей евклидовой метрикой. Эта форма обычно обозначается так: $Edu^2 + 2Fdudv + Gdv^2$, где $E = \langle r_u, r_u \rangle$, $F = \langle r_u, r_v \rangle$, $G = \langle r_v, r_v \rangle$. Обозначим матрицы этих форм соответственно через $Q = \begin{pmatrix} L & M \\ M & N \end{pmatrix}$

и $A = \begin{pmatrix} E & F \\ F & G \end{pmatrix}$. Тогда (см. [1, 5]) следующие две скалярные функции: $K = \det A^{-1}Q$ и $H = \text{Spur } A^{-1}Q$ являются важными геометрическими характеристиками поверхности и называются гауссовой и средней кривизнами соответственно. В данный момент особый интерес представляет для нас средняя кривизна, поскольку именно она тесно связана с геометрией границ раздела сред. Из определения средней кривизны следует, что $H = \frac{GL - 2MF + EN}{EG - F^2}$. Оказывается, выбором подходящих локаль-

ных координат это выражение можно существенно упростить, что позволит связать среднюю кривизну с поверхностным натяжением на границе раздела сред. Напомним, что криволинейные регулярные координаты u, v на поверхности M называются конформными (или изотермическими), если в них индуцированная риманова метрика $ds^2 = Edu^2 + 2Fdudv + Gdv^2$ является диагональной, т. е. $E = G$, $F = 0$ и $ds^2 = E(du^2 + dv^2)$. Оказывается, если метрика на поверхности является вещественно-аналитической, то в окрестности каждой точки всегда существуют локальные конформные координаты [1, 5]. Другими словами, для любой точки $P \in M$ существует окрестность $U(P)$ такая, что в ней можно ввести координаты p и q (являющиеся вещественно-аналитическими функциями от исходных координат), в которых метрика ds^2 имеет вид $\lambda(p, q)(dp^2 + dq^2)$. Можно считать, что поверхности раздела сред являются вещественно-аналитическими во всех внутренних неособых точках, поэтому мы можем без ограничения общности предположить существование в окрестности каждой регулярной точки конформных координат.

Итак, пусть u и v — конформные координаты в некоторой окрестности точки P на поверхности M в $\mathbf{R}^3$. Можно считать, что матрица первой формы является единичной в точке P . В одной точке этого всегда можно добиться соответствующим ортогональным преобразованием конформных координат. В частности, в точке P имеем: $EG - F^2 = 1$. В силу конформности координат мы получаем: $E = G$, $F = 0$ в окрестности точки P , следовательно $H = (L + N)/E$. Значение средней кривизны в точке P выглядит еще проще: $H(P) = L + N$, так как $E(P) = 1$. Вспомним явный вид коэффициентов L и N . Получаем, что $H = E^{-1}(\langle r_{uu}, n \rangle + \langle r_{vv}, n \rangle) = E^{-1}\langle r_{uu} + r_{vv}, n \rangle =$
 $= E^{-1}\langle \Delta r, n \rangle$, где $\Delta = \frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2}$ — оператор Лапласа. Еще

более упростим выражение для средней кривизны. Рассмотрим касательную плоскость к поверхности в точке P , выберем в ней декартовы координаты x, y (рис. 9), а третью координату z направим по нормали к поверхности. Тогда локально поверхность около точки P записывается при помощи радиус-вектора $r(x, y) = (x, y, f(x, y))$, где $f(x, y)$ — гладкая функция, график которой и задает поверхность. Тогда получаем $H(P) = \langle r_{xx} + r_{yy}, n \rangle$. Отметим, что координаты x, y не обязаны быть конформными на поверхности, но они ортогональны в точке P и потому $E(P) = G(P) = 1$, $F(P) = 0$. Так как относительно координат x, y, z вектор нормали $n(P)$ записывается в виде $n(P) = (0, 0, 1)$, то $H(P) = \Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$. Для нас

важно то, что график функции f задает (локально) поверхность M в окрестности точки P . Чтобы связать среднюю кривизну с поверхностным натяжением, вспомним еще одну интерпретацию оператора Лапласа.

Пусть $f(x_1, \dots, x_n)$ — гладкая функция на $\mathbf{R}^n$, отнесенном к декартовым координатам $x_1, \dots, x_n$. Определим ρ -локальное усреднение функции f на $\mathbf{R}^n$ следующим образом: $F_\rho f(x) =$
 $= \frac{1}{\gamma_\rho} \int_{S_\rho} f$, где S_ρ — евклидова сфера в $\mathbf{R}^n$ радиуса ρ с центром

в точке $x = (x_1, \dots, x_n)$, а γ_ρ — объем сферы S_ρ . Другими словами, мы усредняем исходную функцию f по границе шаровой окрестности радиуса ρ с центром в точке x . Если функция задана на плоскости, то S_ρ — это окружность, а γ_ρ — ее длина, $\gamma_\rho = 2\pi\rho$. Определим теперь «отклонение функции от ее ρ -локального усреднения», положив: $\Delta_\rho f(x) = f(x) - F_\rho f(x)$. Переходя к пределу при $\rho \rightarrow 0$, мы можем построить функцию $g(x) = \lim_{\rho \rightarrow 0} \Delta_\rho f(x)$. Легко проверить (предоставляем это читателю),

что функция $g(x)$ совпадает с точностью до постоянного множителя с функцией $\Delta f(x)$. Функцию $\lim_{\rho \rightarrow 0} F_\rho f(x) = Ff(x)$ можно назвать локальным усреднением функции f . Тогда $\Delta f(x) = f(x) - Ff(x)$, т. е. значение оператора Лапласа на функции

f равно отклонению исходной функции f от ее локального усреднения Ff . Такая интерпретация оператора Лапласа полезна со многих точек зрения. Например, если f — гармоническая функция, т. е. $\Delta f = 0$, то она совпадает со своим локальным усреднением. В частности, гармоническая функция не достигает строгого локального максимума ни в какой внутренней точке области, где эта функция определена (максимум достигается на границе).

Возвращаясь к анализу средней кривизны, мы видим, что кривизна $H(P) = f_{xx} + f_{yy} = \Delta f$ может интерпретироваться как отклонение радиус-вектора поверхности

от его локального усреднения в направлении нормали. Напомним, что радиус-вектор полностью определяется функцией $f(x, y)$. Предположим, что поверхность является границей раздела двух сред и что на ней действуют силы притяжения между близкими точками, как это имеет место в случае молекул воды в описанном выше примере. Для наглядности можно считать, что мы имеем дело с мыльной пленкой. Тогда можно выписать цепочку равенств:

давление, которому подвергается поверхность в точке P , равно проекции на нормаль к поверхности равнодействующей всех локальных сил притяжения между точками, близкими к точке P , что в свою очередь равно (с точностью до некоторого постоянного множителя, не зависящего от точки) отклонению радиус-вектора от его локального усреднения (отклонение в направлении нормали). Итак, давление в точке P равно некоторой постоянной, умноженной на среднюю кривизну поверхности в данной точке. Постоянная, возникшая в результате этого анализа, т. е. коэффициент перед средней кривизной, называется поверхностным натяжением. Отсюда и следует анонсированный выше результат. А именно, если двумерная поверхность является мыльным пузырем или системой таких пузырей, или же границей между двумя жидкостями одинаковой плотности (так что силой тяжести можно пренебречь), или мыльной пленкой, натянутой на проволоочный контур, то, при условии, что поверхность находится в равновесии, давление на каждую сторону поверхности является постоянной функцией, т. е. не зависит от точки. Следовательно, средняя кривизна постоянна, что и требовалось.

Этим объясняется сферическая форма мыльных пузырей, которую они приобретают во время свободного падения. В этом случае силой тяжести можно пренебречь. Здесь давление газа внутри сферы превышает наружное давление и равновесие пленки достигается в результате действия сил поверхностного натяжения, стабилизирующих сферическую пленку. Особый интерес представляет случай мыльной пленки, натяну-

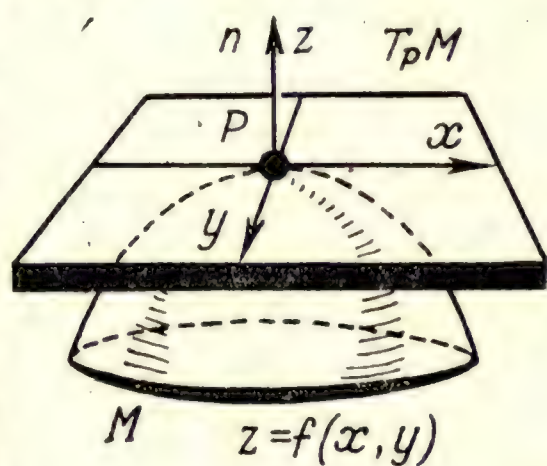


Рис. 9

той на проволочный контур. Здесь разности давлений с обеих сторон пленки нет. Следовательно, давление на одну сторону пленки совпадает с давлением на противоположную ее сторону в окрестности каждой точки на поверхности. Итак, равнодействующая сил равна нулю (силой тяжести мы пренебрегаем), а потому средняя кривизна равна нулю.

5. Мыльные пленки постоянной положительной кривизны и постоянной нулевой кривизны.

Обнаружив, что мыльные пленки — «жидкие поверхности» — являются поверхностями постоянной средней кривизны, Пуассон, естественно, поставил вопрос о полном описании таких поверхностей. Поскольку задача о возможной форме капли жидкости или пузырьков еще не была окончательно решена, то внимание исследователей первоначально сконцентрировалось на случае положительной средней кривизны. Ясно, что стандартная евклидова сфера является такой поверхностью, а потому свободно падающие мыльные пузыри действительно приобретают форму сферы. Однако оставалось неясным — нет ли других замкнутых (т. е. без края) поверхностей постоянной положительной средней кривизны. Конечно, многочисленные физические опыты с мыльными пузырями быстро убеждают нас, что сфера — единственно возможная поверхность постоянной положительной кривизны в классе гладких замкнутых поверхностей, но математическое доказательство этого факта требует некоторых усилий. Пуассон доказал, что в классе сфероидов, достаточно близких к сфере и являющихся поверхностями вращения (т. е. слегка сжатых с полюсов или слегка вытянутых вдоль какой-то оси), сфера является единственной поверхностью с постоянной средней кривизной. Эта информация была практически единственной вплоть до 1853 г., когда Жиллет (Jellet) доказал, что среди замкнутых двумерных поверхностей, имеющих звездчатую форму (т. е. каждая точка которых видна из некоторой внутренней точки, «центра» поверхности), сфера является единственной поверхностью с постоянной положительной средней кривизной. Наконец, уже в нашем столетии было доказано, что компактная двумерная замкнутая поверхность постоянной положительной кривизны, не имеющая самопересечений, является стандартной сферой. Отметим, что мы говорим о поверхностях, вложенных или погруженных в $\mathbb{R}^3$. В частности, вторая форма и средняя кривизна определены лишь для поверхностей, погруженных в некоторое объемлющее пространство.

Итак, мы выделили два основных случая: если мыльная пленка ограничивает замкнутый объем, давление в котором превышает наружное давление, то поверхность имеет постоянную положительную среднюю кривизну; если же мыльная пленка не содержит в себе замкнутых объемов и натянута на проволочный контур, то разности давлений нет и средняя кри-

визна равна нулю. Первые основные эксперименты, выполненные Плато, были посвящены изучению границы раздела жидкостей. При этом опыты проводились с несмешиваемыми жидкостями равной плотности; одна жидкость погружалась в другую, что приводило к появлению капель, ограниченных поверхностями постоянной положительной кривизны. Эксперименты с поверхностями нулевой кривизны начались позже.

В дальнейшем основное внимание мы сосредоточим на случае мыльных пленок с границей. Эти пленки допускают, таким образом, следующие три эквивалентных описания: а) поверхности нулевой средней кривизны; б) поверхности минимальной площади, минимальные поверхности; в) поверхности, в каждой внутренней регулярной точке которых главные кривизны равны по модулю и имеют разные знаки. Последняя характеристика следует из того, что $H = \lambda_1 + \lambda_2$, где λ_1 и λ_2 — главные кривизны.

6. Устойчивые и неустойчивые поверхности.

Уже простейшие эксперименты показывают, что различные поверхности раздела двух сред по-разному реагируют на малые возмущения. Некоторые из них сопротивляются разрушению и оказываются устойчивыми, стабильными. Другие сразу же разрушаются, оказываются нестабильными, неустойчивыми. Здесь проявляется важный эффект, наглядно иллюстрируемый уже на уровне скалярных функций на гладком многообразии. Напомним, что если $f(x)$ — гладкая функция на конечномерном многообразии M , то большую роль в изучении поведения функции играют критические точки, т. е. такие, в которых $\text{grad } f = 0$. Эти точки могут быть нескольких сортов. Во-первых, критическими являются точки локальных минимумов и максимумов функции. Однако кроме них есть еще так называемые седловые точки. Типичный вид такой точки условно изображен на рис. 10. Они характеризуются наличием двух сортов направлений: вдоль направлений a (рис. 10) функция строго увеличивается, а вдоль ортогональных направлений b — строго уменьшается. Количество независимых направлений, вдоль которых функция строго уменьшается, называется обычно индексом критической точки. Максимальный индекс имеют точки невырожденных локальных максимумов. Совершенно аналогичная картина наблюдается и в том случае, когда функция (функционал) определена на бесконечномерном пространстве. Таким пространством является (при правильном выборе формулировки, которую мы здесь не будем уточнять) множество всех двумерных поверхностей, например, затягивающих проволочные контуры в трехмерном пространстве. Тогда функционал площади сопоставляет каждой такой поверхности ее площадь. Изменение, деформация поверхности порождает соответствующее изменение значения функционала площади. Если двумерная пленка находится в равновесии, это означает, что, рассматриваемая как точка бесконечномерного

пространства, она является критической для функционала площади. Но, как и в конечномерном случае, это положение равновесия может быть нескольких сортов: локальный минимум, локальный максимум и седло. В последнем случае существуют достаточно малые возмущения пленки, которые уменьшают ее площадь. При физической реализации устойчивых пленок мы получаем, что в случае локального минимума функционала площади пленка устойчива, стабильна, т. е. сопротив-

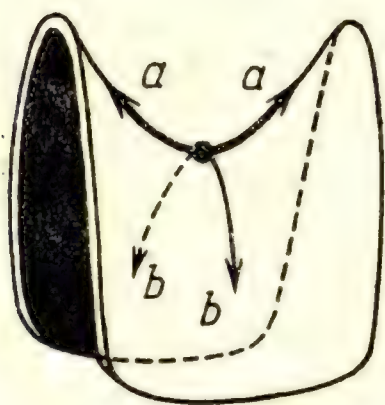


Рис. 10

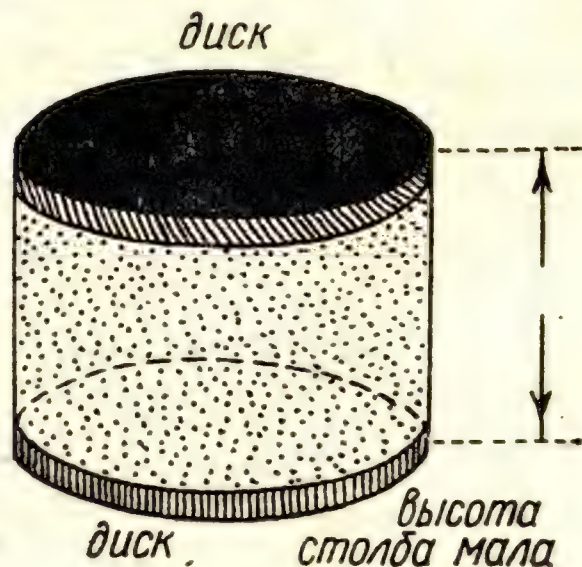


Рис. 11

ляется малым возмущениям и стремится вернуться в исходное равновесное состояние, а в случае седла пленка неустойчива, нестабильна и существуют возмущения, сколь угодно малые по амплитуде, которые уменьшают площадь пленки, и она начинает быстро деформироваться, приобретая новую конфигурацию, отвечающую меньшему значению энергии. При этом новая конфигурация может существенно отличаться от исходной даже с топологической точки зрения (см. примеры ниже). «Бесконечномерность» пространства всех поверхностей означает также, что существует бесконечно много различных независимых направлений возмущений пленки.

7. Опыты Плато по устойчивости столба жидкости.

Плато провел серию экспериментов с жидкостями. О некоторых из них мы уже говорили. Здесь мы вкратце опишем следующую серию его опытов. Между двумя металлическими дисками одинакового радиуса, центры которых расположены на одной прямой, ортогональной плоскости дисков, Плато реализовывал столбик жидкости, который принимал форму прямого кругового цилиндра, целиком состоящего из жидкости (рис. 11). Поскольку граница цилиндра является поверхностью постоянной положительной средней кривизны, то теоретически такой столб жидкости является критической точкой в пространстве всех «поверхностей». Не следует смешивать этот случай со случаем замкнутых (!) компактных гладких поверхностей постоянной положительной средней кривизны (как мы уже знаем, такая поверхность является сферой). Цилиндр не является компактной поверхностью, если он продол-

жен до бесконечности в обе стороны. Если же он ограничен двумя краями — окружностями (дисками), то он не является замкнутым (т. е. без края) гладким многообразием. Свойства цилиндрического столба жидкости существенно зависят от его высоты. В реальных физических опытах можно реализовать

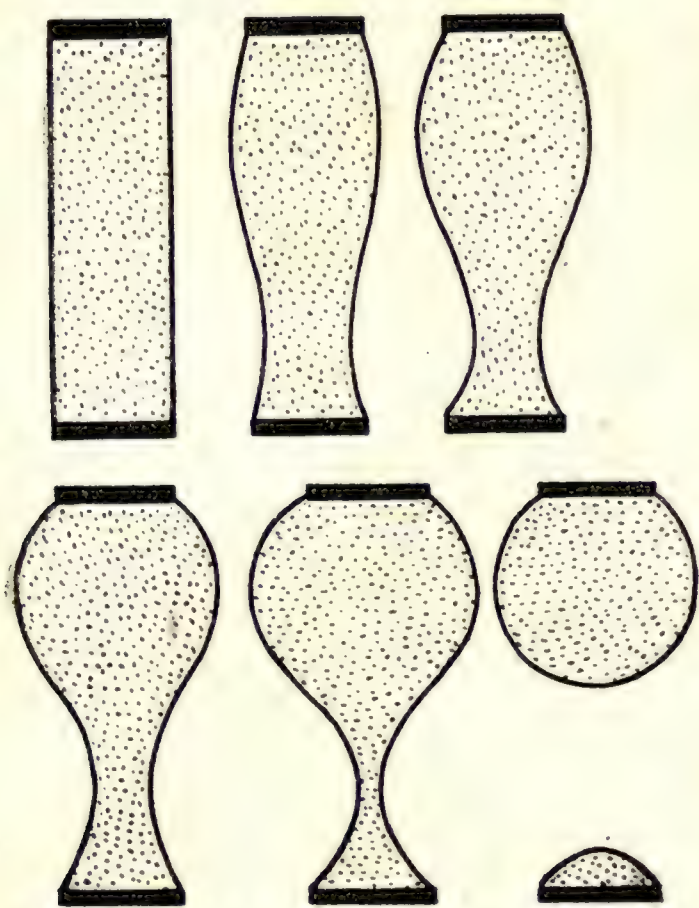


Рис. 12

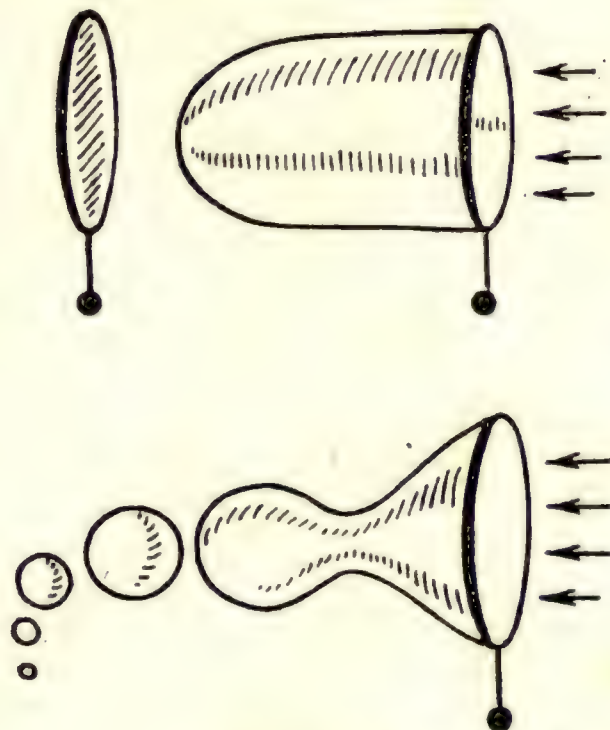


Рис. 13

«не очень высокий» столб жидкости между двумя параллельными дисками. Если два диска значительно удалены друг от друга, то столб жидкости разрушается. Крайне интересно наблюдать за процессом разрушения этого столба, когда два диска, оставаясь параллельными, постепенно удаляются друг от друга, растягивая столб жидкости по вертикали. Оказывается, водяной столб устойчив лишь в том случае, когда его высота не превышает приблизительно утроенного диаметра граничного диска. Если же высота столба приближается к утроенному диаметру (приблизительно 3,1 диаметра) и затем превышает эту величину, то столб начинает разрушаться. Если высоту столба увеличивать крайне медленно и осторожно, то поведение столба вблизи критического значения его высоты можно проследить визуально. Его деформация происходит в этом случае достаточно медленно, что позволяет изобразить процесс перестройки столба с помощью рис. 12. Таким образом, невысокие цилиндры устойчивы, а излишне высокие — неустойчивы. Критической высотой является приблизительно утроенный диаметр диска. Конечно, эти эксперименты проводятся с дисками достаточно малого диаметра, чтобы силы поверхностного натяжения могли успешно конкурировать с силой тяжести, действующей на столб жидкости.

Процесс распада неустойчивого столба, т. е. процесс его

качественной перестройки, можно продемонстрировать также в другом реальном опыте, проведенном еще в XVII в. Рассмотрим проволочный контур в виде правильной плоской окружности и, опустив его в мыльную воду, натянем на него диск — мыльную пленку. Затем начнем быстро перемещать контур в направлении, ортогональном его плоскости, т. е. ортогонально к мыльной пленке (рис. 13). Под воздействием набегающей струи воздуха мыльная пленка прогибается и сначала приобретает приблизительно форму прямого кругового цилиндра (во всяком случае, это происходит в непосредственной близости от контура). Однако затем эта цилиндрическая форма быстро разрушается, на ней появляются перетяжки, пленка подвергается перестройке и от нее начинают отрываться мыльные пузыри, имеющие практически сферическую форму. Этот же процесс, очевидно, можно проследить, фиксируя контур в пространстве и направив на него струю воздуха, ортогональную плоскости контура. Цилиндрический столб жидкости был первой критической поверхностью, устойчивость которой была аналитически доказана Плато. Тогда Бир (Beer) предположил, что все поверхности постоянной средней кривизны являются локальными минимумами функционала площади. Плато доказал, что это предположение неверно. Он рассмотрел цилиндр высоты большей, чем критическая. С математической точки зрения такой цилиндр по-прежнему является критической точкой (поверхностью) для функционала площади в пространстве всех поверхностей. Плато сначала аналитически доказал, что такой цилиндр по-прежнему имеет локально наименьшую возможную площадь в специальном классе всех возмущений этого цилиндра, сохраняющих площадь каждого горизонтального сечения. Другими словами, этот класс возмущений характеризуется тем, что отсутствуют перемещения жидкости по вертикали. Перемещения жидкости по горизонтали, сохраняющие площадь любого плоского сечения, ортогонального первоначальной оси цилиндра, допускаются. Другими словами, если мы возмущаем цилиндр так, что он слегка изгибается, но при этом плоские сечения сохраняют свою площадь, то возмущенный цилиндр имеет площадь, большую, чем исходный. Напомним, что если высота цилиндра меньше критической, то любое (!) возмущение цилиндра увеличивает его площадь, т. е. в этом случае исходная поверхность является точкой локального минимума для функционала площади в пространстве всех поверхностей. Если же высота больше критической, то цилиндр является седловой точкой в пространстве всех поверхностей. Причем существует по-прежнему бесконечно много независимых направлений α возмущения этой критической поверхности в пространстве всех поверхностей, увеличивающих его площадь. Бесконечность числа направлений, вдоль которых критическая поверхность является минимумом функционала площади, следует из того,

что в горизонтальных плоских сечениях имеется бесконечно много различных деформаций α , сохраняющих площадь сечений. С другой стороны, если высота цилиндра больше критической, то существует по крайней мере одно направление β возмущения поверхности, вдоль которого площадь начинает уменьшаться. Как мы уже знаем, для создания такого возмущения достаточно придать жидкости движение по вертикали,

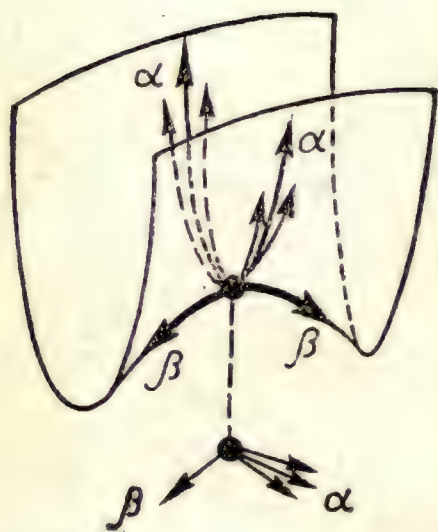


Рис. 14



Рис. 15

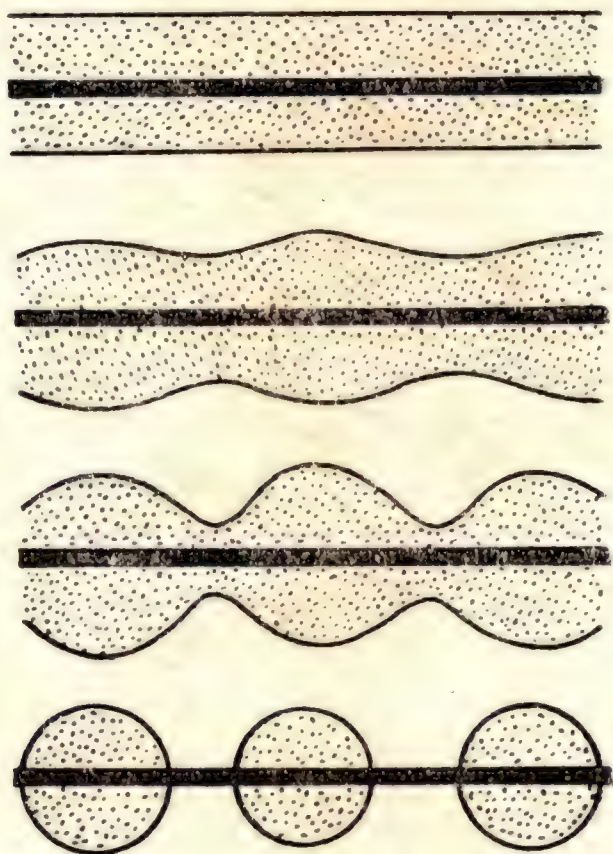


Рис. 16

в результате чего часть цилиндра разбухает, а другая часть, напротив, сужается (рис. 12). Условная схема седла показана на рис. 14. Более того, Плато продемонстрировал на опыте седловой характер критической поверхности в том случае, когда высота цилиндра превышает утроенный диаметр граничного диска. Он реализовал между двумя дисками столб жидкости, высота которого лишь немного больше критической. Столб начал медленно деформироваться в вертикальном направлении, оставаясь при этом симметричной поверхностью вращения. В этот момент тонкой стеклянной палочкой, расположенной горизонтально, наносился легкий удар по столбу жидкости, т. е. физически реализовывалось возмущение, сохраняющее площадь плоских горизонтальных сечений. Столб слегка отклонялся, однако практически мгновенно возвращался в прежнее осесимметричное положение, продолжая при этом медленно деформироваться в вертикальном направлении. Это и указывает на устойчивость поверхности относительно возмущений указанного типа α (рис. 15). Деформация же в вертикальном направлении реализуется возмущением типа β . Интересно, что распад достаточно длинного столба жидкости происходит волнообразно. Для наглядной демонстрации этого процесса можно взять тонкую нить, покрытую жидким цилиндрическим слоем так, что нить является осью этого цилиндра.

Тогда происходит перестройка этого цилиндра, показанная на рис. 16. Цилиндр становится волнообразным, затем каждое вздутие превращается в шар, нанизанный на прежнюю нить-ось. Ясно, что это — суммарный результат действия сил поверхностного натяжения и неустойчивости длинных цилиндров жидкости. Конечно, изображенный на рис. 16 процесс является некоторым упрощением реальной картины, которая существенно сложнее, хотя в результате все равно получается ожерелье шарообразных капель, нанизанных на исходную нить.

Этот процесс изучался особенно подробно, поскольку задача покрытия тонкого проводника (проволоки) равномерным цилиндрическим слоем изоляции возникла (и была решена) в электротехнике. При этом важно было научиться стабилизировать длинный цилиндрический слой, окутывающий проволоку, идущую по оси цилиндра.

8. Физическая реализация геликоида.

В 1842 г. Каталан (Catalan) доказал, что единственными полными линейчатыми поверхностями нулевой средней кри-

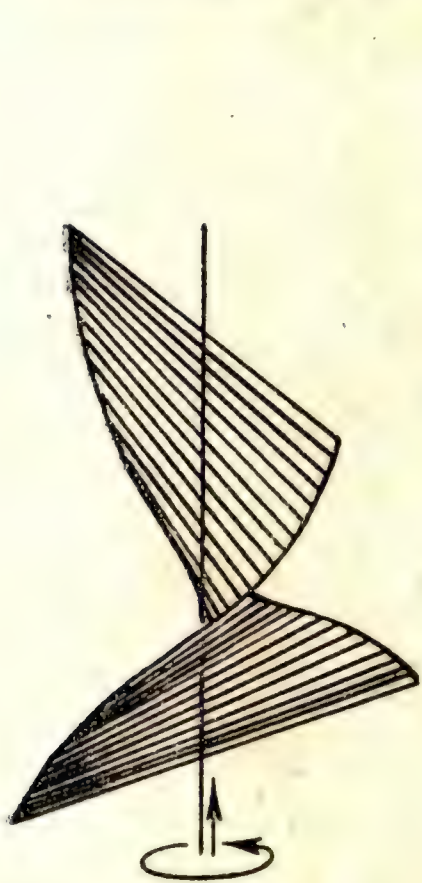


Рис. 17

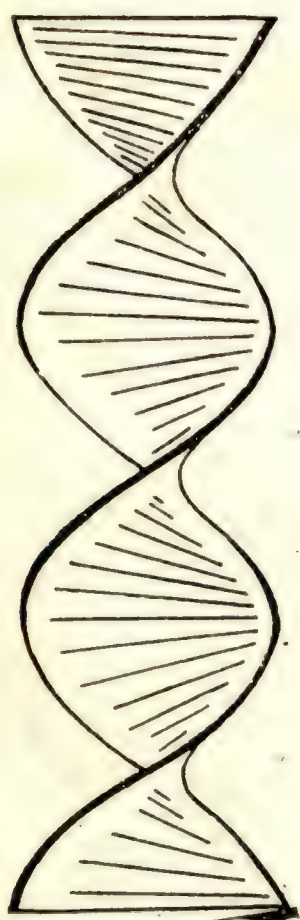


Рис. 18

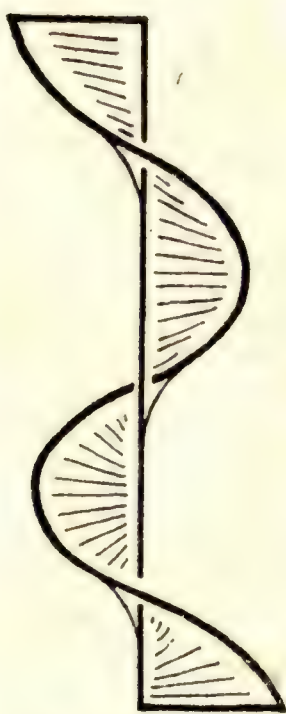


Рис. 19



Рис. 20

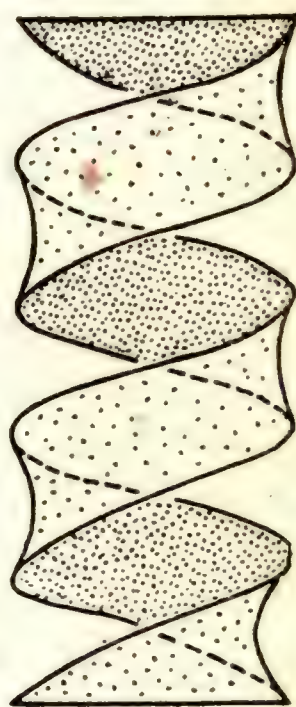


Рис. 21

визны являются плоскость и геликоид. Геликоид получается в результате композиции двух движений прямой линии: поступательного с постоянной скоростью и вращательного с постоянной угловой скоростью в плоскости, ортогональной вектору трансляции (рис. 17). Другими словами, следует взять прямую, пересекающую вторую прямую под прямым углом, и равномерно перемещать первую прямую вдоль второй, одновременно равномерно вращая ее. Эта поверхность некомпактна. Если ограничиться лишь отрезком вместо целой прямой, то получится поверхность, показанная на рис. 18. В координатах (x, y, z) и (r, t) она задается следующим радиус-вектором: $x = r \cos \omega t$, $y = r \sin \omega t$, $z = at$, где a и ω — скорости поста-

тельного и вращательного движений соответственно. Плато реализовал «половину» геликоида, используя для этого спираль из проволоки, обвивающую осевую прямую линию (рис. 19). Однако попытка реализовать полный геликоид с помощью пленки, натянутой на проволочный контур, наталкивается на трудности. Конечно, можно получить часть полного геликоида, составив вместе два экземпляра пленки, изображенной на рис. 19. Однако при таком подходе мы получаем, что кроме граничного контура, составленного из двух спиралей с общей осью, для реализации мыльной пленки следует добавить и саму эту ось, например, в виде тонкой нити, идущей по оси всей конструкции (рис. 20). Эта нить стабилизирует пленку и препятствует ее перестройке в минимальную поверхность, уже не являющуюся геликоидом. В самом деле, если убрать эту осевую нить и оставить только две граничные спирали, то в том случае, когда шаг спиралей не очень велик, т. е. когда спираль делает много витков на сравнительно небольшом расстоянии вдоль оси, геликоид превращается в поверхность, показанную на рис. 21. Ясно, что эта поверхность не является линейчатой, т. е. не состоит из отрезков прямых. Попытки реализовать часть прямого геликоида в виде мыльной пленки, затягивающей замкнутый проволочный контур в виде двух соосных спиралей с малым шагом, заканчиваются неудачей. При этом предполагалось, что контур не должен содержать никаких дополнительных проволочных отрезков, расположенных строго внутри пленки и не являющихся краем пленки. Здесь проявляется существенное отличие прямого геликоида от поверхностей, с которыми мы знакомились ранее. Дело в том, что геликоид — некомпактная поверхность, существующая (с математической точки зрения) независимо от какого-либо граничного контура. В реальных же опытах мы можем реализовывать только мыльные пленки, затягивающие те или иные компактные проволочные контуры. Накладывая это дополнительное условие на реализацию геликоида, мы и сталкиваемся с отмеченными выше трудностями. Если отказаться от попыток реализации геликоида с малым шагом и поставить перед собой задачу реализации геликоида с достаточно большим шагом граничных спиралей, то построить такую мыльную пленку не составляет труда. Для этого нужно рассмотреть контур, составленный из двух соосных спиралей. Ясно, что при достаточно большом шаге этот контур ограничивает мыльную пленку — геликоид (рис. 22). Доказательство линейчатой структуры этой пленки можно дать следующим наглядным образом. Ясно, что геликоид на рис. 22 получается из минимальной поверхности на рис. 21 растяжением граничных спиралей в направлении оси. С увеличением шага спирали мыльная пленка деформируется. Фиксируем в пространстве ось спиралей и направление нашего взгляда на пленку. Будем следить за деформацией части пленки h , расположенной в непосред-

ственной близости от точки скрещивания спиралей (рис. 23). Эта часть пленки имеет в проекции на сетчатку глаза вид криволинейного треугольника, который, очевидно, уменьшается по мере растяжения спиралей. В некоторый момент времени этот треугольник исчезнет, превратившись в точку, и мы получим

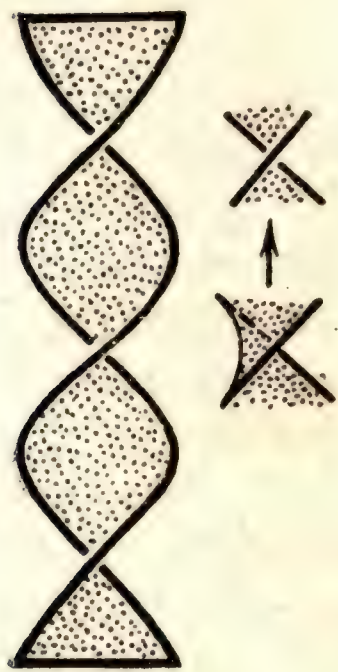


Рис. 22

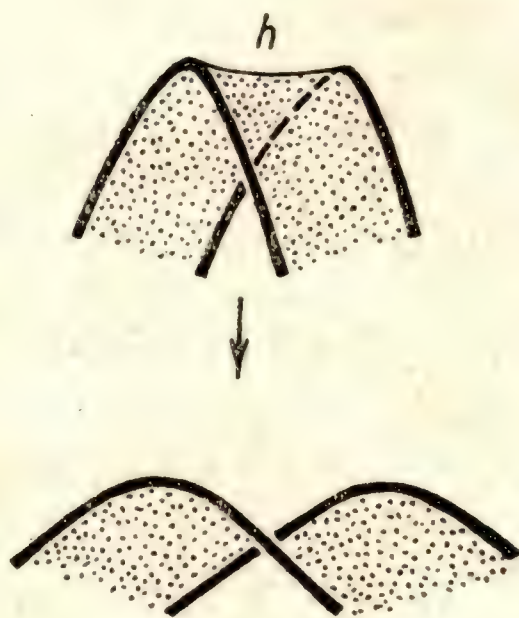


Рис. 23

пленку, показанную на рис. 22. Так как проекция пленки имеет теперь вид, показанный на рис. 23, то при вращении пленки вокруг ее оси спиралей начинают бежать, например, слева направо. В этом же направлении движутся и точки скрещивания спиралей. Так как часть пленки, проектирующаяся (на сетчатке глаза) в каждую точку скрещивания, является прямолинейным отрезком, соединяющим точку на ближайшей спирали с точкой на дальней спирали, то вся пленка составлена из прямолинейных отрезков, ортогональных оси, что и дока-

зывает линейчатость пленки и ее совпадение с геликоидом.

9. Физическая реализация катеноида и его перестройки при изменении граничного контура.

Следующий популярный пример минимальной поверхности получается так. Рассмотрим минимальную поверхность вращения, образованную вращением вокруг оси Ox кривой, задаваемой уравнением $y = a \cdot \operatorname{ch} \frac{x}{a}$, где a — ненулевая постоянная. Хорошо известно, что эта кривая совпадает с формой провисания тяжелой цепи, закрепленной в двух точках (рис. 24). Сила тяжести направлена здесь вниз по оси Oy . Прямое вычисление показывает, что средняя кривизна получающейся поверхности вращения равна нулю: $H = \lambda_1 + \lambda_2 = \frac{1}{y \sqrt{1 + (y')^2}} - \frac{y''}{(1 + (y')^2)^{3/2}} = 0$. Таким образом, мы получили минимальную поверхность, называемую катеноидом. Ее можно рассматривать как мыльную пленку, затагивающую две соосные окружности, расположенные в параллельных плоскостях. Если мы перепишем уравнение цепной линии следующим образом: $\frac{y}{a} = \operatorname{ch} \frac{x}{a}$, то получаем, что все цепные линии подобны и, следовательно, все катеноиды эквивалентны в том смысле, что могут быть переведены друг в друга движением

трехмерного пространства и изменением масштаба. В этом смысле цепные линии похожи на параболы, которые также эквивалентны относительно движений плоскости и изменения масштаба (в отличие от гипербол и эллипсов). Катеноид получается вращением цепной линии лишь в том случае, когда эта линия находится на определенном расстоянии от оси вращения. Если цепную линию сместить с этого положения, то получающаяся поверхность вращения уже не является, вообще говоря, минимальной. Так, например, если цепная линия с малым провисанием почти касается оси Ox , то при ее вращении низшая точка линии, ближайшая к прямой, описывает окружность малого радиуса. Следовательно, один из радиусов

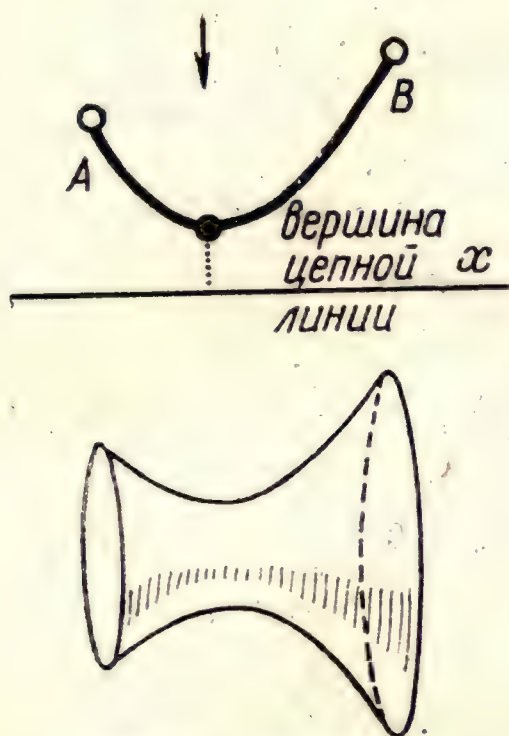


Рис. 24

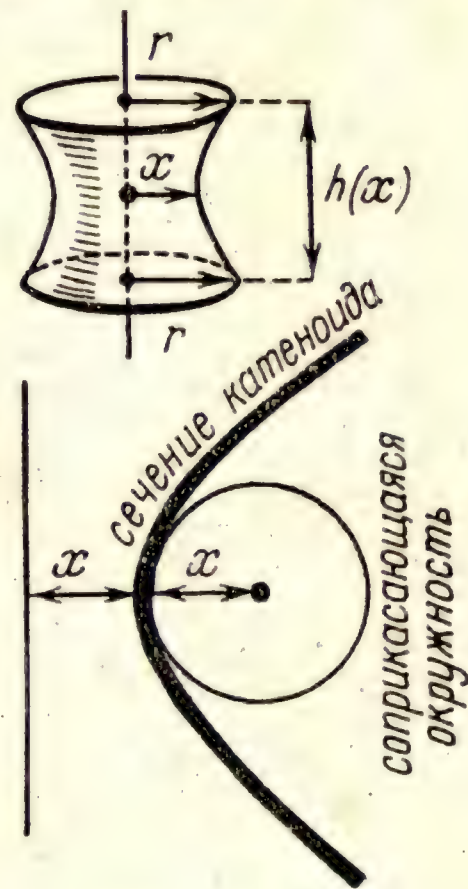


Рис. 25

кривизны очень мал, а второй велик, следовательно, сумма главных кривизн в точках, расположенных на горловине поверхности вращения, отлична от нуля.

Рассмотрим зависимость структуры катеноидов от размера граничного контура. Рассмотрим для простоты две соосные окружности одинакового радиуса r и пусть h — расстояние между содержащими их параллельными плоскостями (рис. 25). Будем считать, что радиус фиксирован. Введем еще один параметр x — радиус горловины катеноида, т. е. расстояние до оси вращения от ближайших к ней точек катеноида, образующих окружность радиуса x . Ясно, что параметры x и h связаны. Можно считать, что h — гладкая функция от x . Опишем качественные характеристики функции $h(x)$. Характер зависимости h от x определяется уже отмеченным выше обстоятельством: так как средняя кривизна катеноида равна нулю, то радиус кривизны — радиус соприкасающейся окружности в вершине цепной линии (в ближайшей к оси вращения точке) — равен радиусу окружности, описываемой вершиной линии при ее вращении вокруг оси, т. е. равен x . Следова-

но, чем меньше x , тем больше кривизна цепной линии в ее вершине. Параметр x изменяется на отрезке от нуля до бесконечности. Если $x > r$, то очевидно, что никакой минимальной поверхности, натянутой на две окружности, не существует (рис. 26). На рисунке жирной линией изображено сечение катеноида вертикальной плоскостью, проходящей через ось симметрии, а тонкой линией — соприкасающаяся окружность в вершине цепной линии. При $x = r$ очевидно, что $h = 0$, т. е. две окружности соприкасаются. При $x < r$ расстояние h начинает

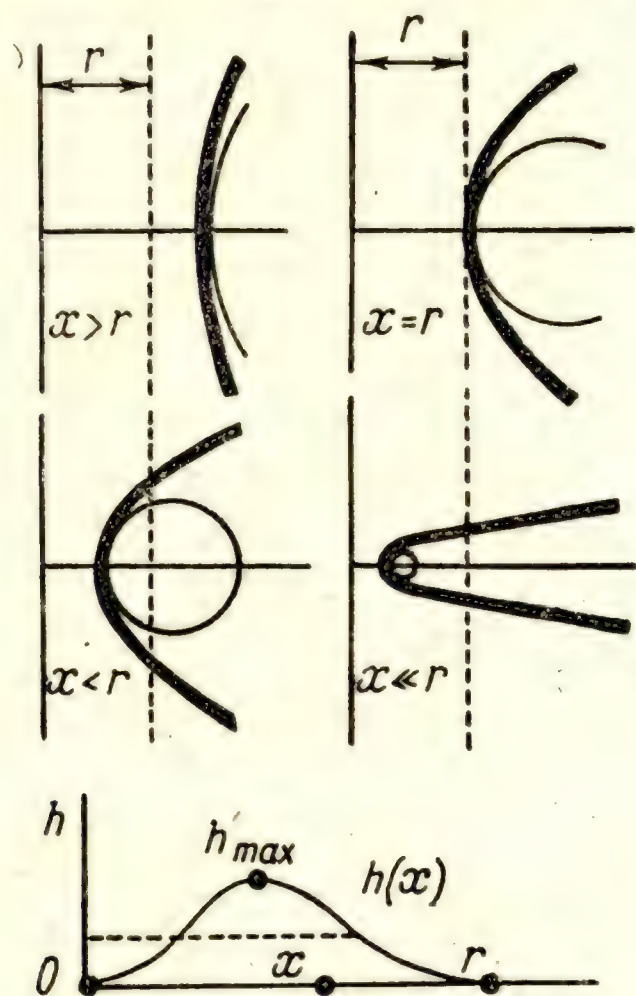


Рис. 26

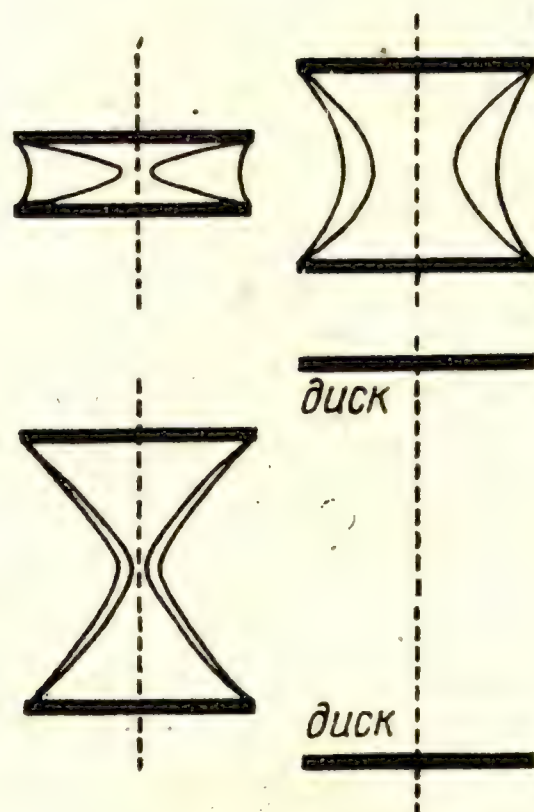


Рис. 27

увеличиваться, окружности раздвигаются, и на них повисает минимальная поверхность. При дальнейшем уменьшении x расстояние h продолжает некоторое время увеличиваться, затем, как это очевидно из рис. 26, достигнув некоторого максимального значения $h_{\max}$, начинает уменьшаться. Когда x стремится к нулю, h также стремится к нулю, цепная линия все ближе приближается к оси Ox . Получившийся график функции $h(x)$ см. на рис. 26. Когда x близко к нулю, то радиус кривизны в вершине цепной линии также мал, что и заставляет эту линию приближаться к горизонтальной оси. Максимальное значение $h_{\max}$ приблизительно равно $\frac{4}{3}r$. Когда

расстояние h между окружностями превышает $h_{\max}$, то катеноид разрывается и минимальная поверхность превращается в два плоских диска, затягивающих граничные окружности. В этом случае никаких других мыльных пленок с этой границей нет. Описанный переход от одного решения к другому происходит скачком, изменяющим топологию пленки. Полезно рассмотреть эволюцию катеноида при увеличении h от нуля до $h_{\max}$. На рис. 27 показано несколько последовательных по-

ложений мыльной пленки. Любопытным фактом является наличие при любом h , $0 < h < h_{\max}$, двух мыльных пленок — катеноидов: внешнего и внутреннего. При увеличении h эти две пленки начинают провисать по направлению к оси симметрии и одновременно сближаться. Наконец, когда h превышает критическое значение, обе пленки лопаются и перестраиваются в пару дисков, затягивающих окружности. Итак, для любого значения h , заключенного в интервале $0 < h < h_{\max}$, существуют три мыльные пленки с данной границей — два катеноида и пара дисков. Плато предположил, что внутренний катеноид неустойчив в отличие от внешнего катеноида и обоих дисков, но не смог это доказать. Попытки реализовать в виде физической мыльной пленки внутренний катеноид заканчиваются неудачей. Предположение о неустойчивости внутренней пленки оказывается справедливым.

10. Изменение топологического типа минимальных поверхностей в зависимости от их устойчивости или неустойчивости.

На один и тот же контур можно натянуть, вообще говоря, несколько различных минимальных поверхностей. Некоторые такие случаи нам уже знакомы, например, два катеноида и

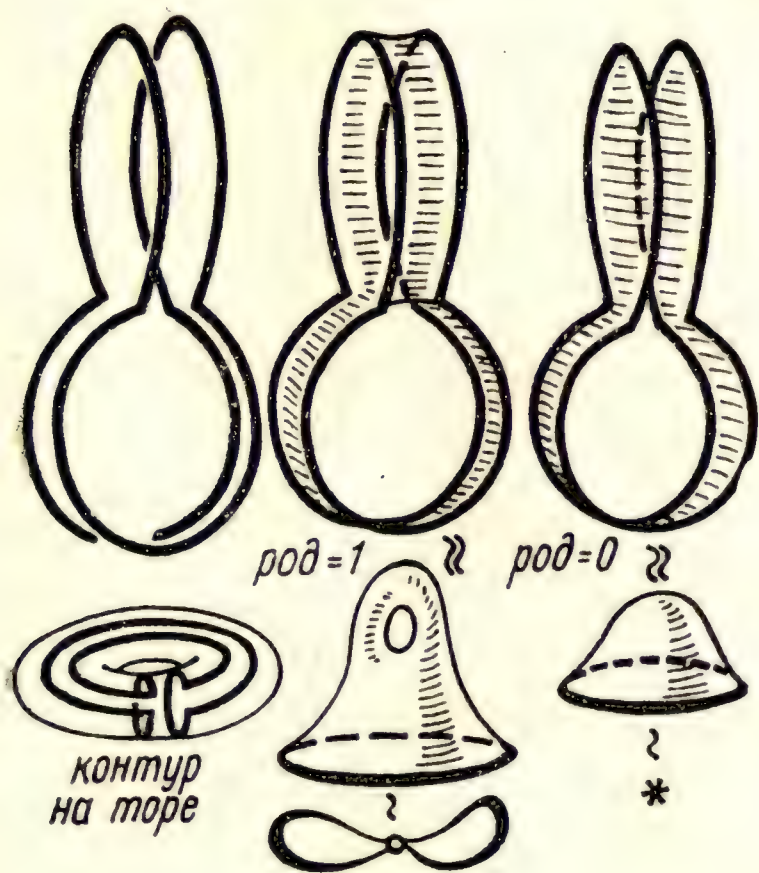


Рис. 28

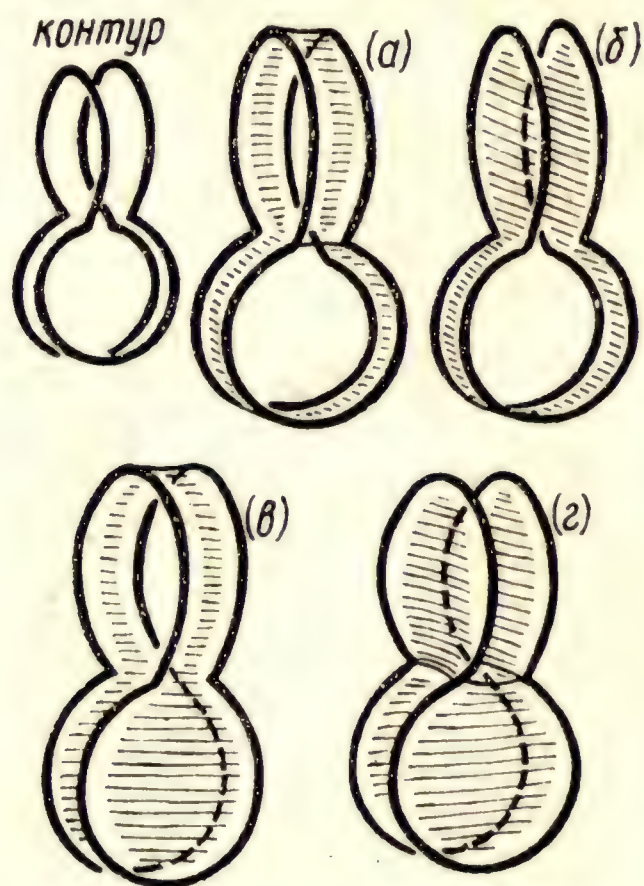


Рис. 29

пара дисков. Единственность мыльной пленки можно гарантировать, например, для плоских замкнутых несамопересекающихся контуров, т. е. реализующихся в виде системы кривых, вложенных в одну плоскость. Ясно, что любая пленка, отличная от части плоскости, ограниченной этими кривыми, имеет большую площадь, чем плоская пленка. При этом можно рассматривать две вариационные задачи: отыскание пленок наименьшей площади при заданном контуре и отыскание минимальных пленок наименьшего топологического рода. При этом

не исключено, что пленка, имеющая наименьшую возможную площадь, имеет тем не менее не минимальный топологический род. И напротив, минимальная пленка наименьшего рода может не реализовывать минимум площади в классе всех пленок с данной границей. Другими словами, абсолютный минимум функционала площади может достигаться на пленках не минимального топологического рода (рис. 28). Если изображенные пары окружностей достаточно близки, то площадь первой пленки меньше площади второй. В то же время первая пленка гомеоморфна тору с дыркой (рис. 28), т. е. имеет в этом смысле род 1, а вторая пленка гомеоморфна диску и имеет род ноль. Поэтому мы будем пока говорить о минимальных пленках, понимая под этим их локальную минимальность и не ставя в полном объеме вопрос об их глобальной минимальности и о связи абсолютных минимумов функционала площади с топологическим родом пленки.

Рассмотрим контур на рис. 28. Оказывается, он достаточно удобен для того, чтобы проследить за нетривиальными перестройками минимальных поверхностей, за изменением топологического типа в зависимости от устойчивости или неустойчивости мыльных пленок. Ясно, что контур гомеоморфен окружности. Будем считать, что он реализован в $\mathbb{R}^3$ в виде проводящего замкнутого контура. Пусть u и v — расстояния между верхними кольцами и нижними кольцами (рис. 28). При каждом фиксированном состоянии контура существует, вообще говоря, несколько типов минимальных поверхностей. Наша цель — исследовать их перестройки при деформации контура. Так как формальное аналитическое исследование минимальных поверхностей в целом затруднительно, мы были вынуждены обратиться к физическому эксперименту. Ниже излагаются результаты анализа, проведенного автором совместно со студентом А. Тужилиным. В опытах были получены следующие мыльные пленки: (а), (б), (в), (г), (д) (рис. 29—31), а также пленки (е), (ж), (з) на рис. 32. Пленки (б), (в), (г), (д) диффеоморфны диску. Пленки типа (е) уже не являются многообразиями и стягиваются по себе в точку (гомотопически тривиальны), как и пленки типов (б), (в), (г), (д). Ограничимся пока изучением лишь пленок (б), (в), (г), (д), т. е. пленок, диффеоморфных диску. Т. Постон в [19] указал на следующую перестройку пленок, возникающую при непрерывной деформации контура.

Возьмем пленку (б) и будем деформировать граничный контур, вызывая тем самым деформацию затягивающей его пленки, так, как это показано на рис. 30. Будем раздвигать две нижние окружности, постепенно удаляя их друг от друга. Можно считать, что этот процесс приблизительно совпадает с растяжением катеноида, когда два его края удаляются друг от друга. Здесь мы пользуемся тем, что при достаточно малом начальном расстоянии окружностей друг от друга мыльная

пленка — лента, затягивающая каждую пару окружностей, достаточно хорошо аппроксимируется катеноидом. При растяжении катеноида наступает критический момент, когда катеноид «схлопывается», перестраивается в пару дисков, затягивающих две удаляющиеся друг от друга окружности. Однако в нашем случае эти два диска все еще соединены узкой лентой, натянутой на две верхние окружности, которые в меньшей степени затронуты описанной деформацией. В результате мы получаем пленку, показанную на рис. 30. Сдвигая теперь две нижние окружности в обратном направлении и возвращая их в первоначальное положение, мы уже не меняем топологи-

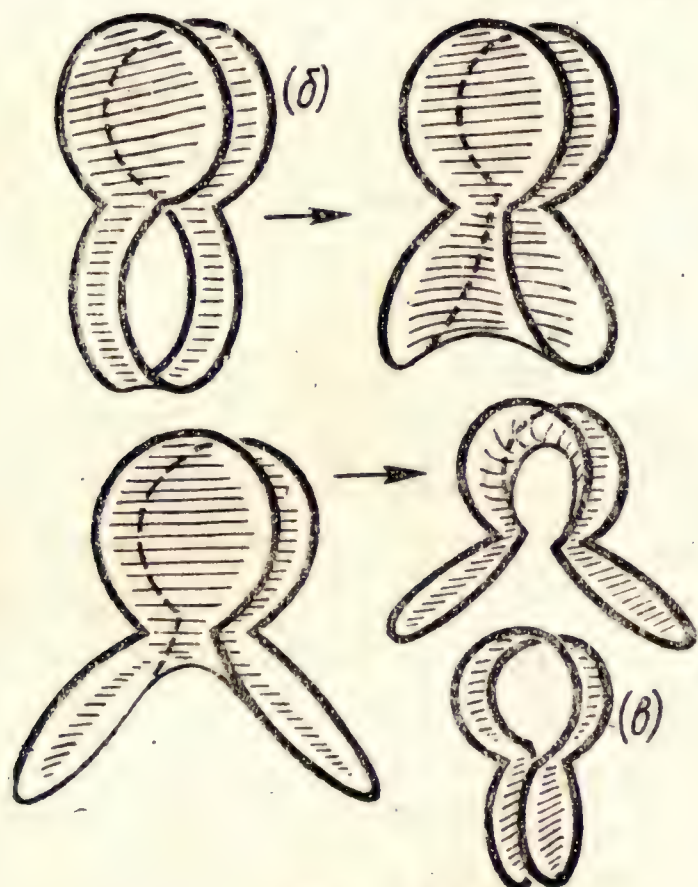


Рис. 30

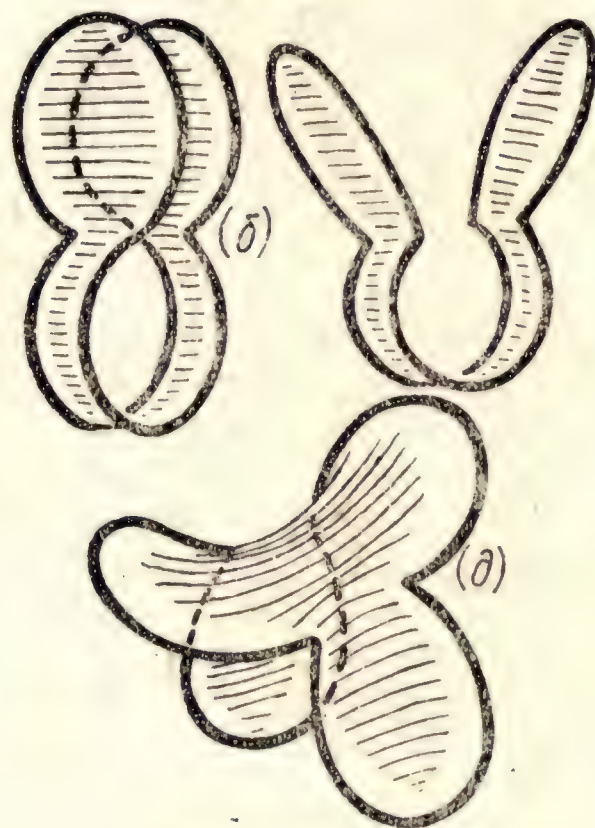


Рис. 31

ского типа пленки и в результате приходим к пленке (в). Итак, мы построили непрерывную деформацию контура, в результате которой произошло скачкообразное изменение пленки, она изменила свой тип: из положения (б) мы перешли в положение (в).

Используя симметрию контура, мы можем аналогичным образом совершить переход в обратном направлении от пленки (в) к пленке (б). Достаточно выполнить описанную выше операцию, но только с двумя верхними окружностями, постепенно удаляя их друг от друга, а затем возвращая их в прежнее положение. В процессе деформации пленка перестроится скачком и пленка (в) превратится в пленку (б). Этот «маятник» можно реализовывать много раз, если мыльная пленка получена из достаточно эластичной жидкости.

Наряду с описанной скачкообразной перестройкой пленки имеется другая замечательная деформация граничного контура, приводящая к плавному изменению пленки с тем же конечным результатом (!). В самом деле, рассмотрим пленку (б) и, в отличие от предыдущего случая, начнем раздвигать верхние окружности так, как это показано на рис. 31. При

этом топологический тип пленки не меняется, она подвергается гомеоморфизму. Затем начнем раздвигать нижние окружности так, как показано на рис. 31. В результате мы получим симметричную пленку седлового типа: ее центр является седловой точкой. По одному направлению пленка поднимается вверх, по другому — опускается. Эта деформация — гладкая, без скачков. Так как пленка (д) симметрична, то, применяя ту же операцию, мы можем гладким образом перейти теперь от нее к пленке (в). Аналогично, повторяя все шаги в обратном по-

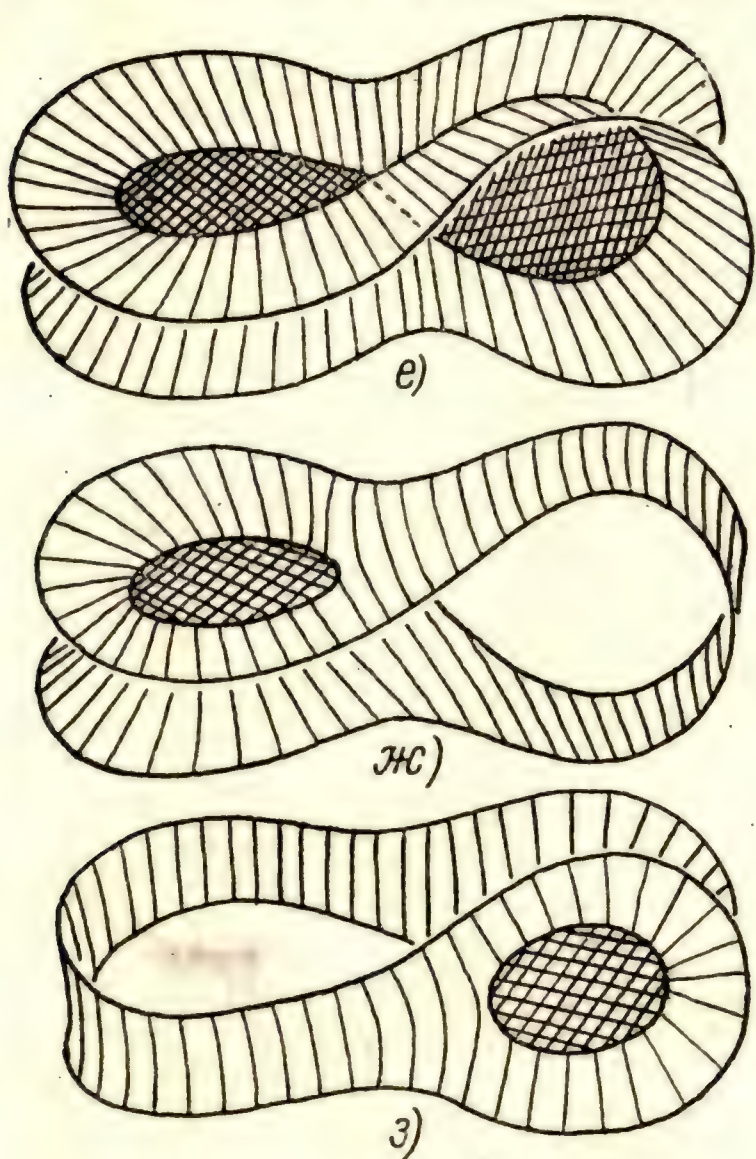


Рис. 32

рядке, мы можем гладко перевести пленку (в) в пленку (б) без скачков и топологических перестроек. Итак, в пространстве всех поверхностей с гомеоморфными граничными контурами указанного типа мы получаем два существенно различных пути, соединяющих положения (б) и (в). Первый путь был реализован гладкой деформацией контура, в процессе которой пленка претерпевала скачкообразную перестройку. Второй путь реализуется другой гладкой деформацией, в процессе которой пленка меняется без скачков, но, как и в первом случае, переходит из положения (б) в (в). В первом приближении можно считать, что описанные деформации контуров можно описать изменением двух вещественных параметров: u — расстояние между двумя верхними окружностями, v — между нижними. Другими словами, u и v — ширина верхнего и нижнего колец-лент на поверхности (а) (рис. 29). Можно считать, что каждое положение граничного контура определяется заданием пары вещественных чисел, которые для наглядности можно считать координатами точки на двумерной плоскости.

Т. Постон в [19] предложил условную схему, предлагая описывать этот процесс с помощью так называемой сборки Уитни. Однако мы решили подойти к этому вопросу с иной точки зрения, а именно — исследовать свойства реального графика площади мыльных пленок. Оказалось, что получающаяся картина отличается от сборки Уитни и описывается ласточкиным хвостом. Уточним постановку задачи. Для каждой точки (u, v) мы имеем соответствующий контур, на который натянуто, вообще говоря, несколько пленок. Вычисляя их площади, получаем график многозначной функции, задающей площадь минимальных поверхностей с данной границей. Пленки бывают

двух сортов: устойчивые и неустойчивые. Среди перечисленных выше пленок неустойчивой является пленка (г); остальные — устойчивы. Рассмотрим более внимательно неустойчивые пленки. Мы будем опираться здесь на один из принципов Плато (см. следующие пункты), согласно которому часть минимальной пленки, ограниченной произвольным замкнутым контуром, нарисованным на пленке, также является минимальной по отношению к этому контуру. В реальных экспериментах неустойчивые пленки можно получать так. Неустойчивость пленки означает, что существуют сколь угодно малые ее возмущения, разрушающие пленку или заставляющие ее существенно изменить положение относительно граничного контура. В то же время, если удачным образом расположить на пленке достаточно много нитей, концы которых прикреплены к контуру, то можно добиться ее стабильности, если сеть нитей достаточно густая. Другими словами, можно пытаться удерживать неустойчивую мыльную пленку в стабильном положении за счет увеличения граничного контура, добавляя к исходному контуру сетку из нитей. Возьмем, например, неустойчивую пленку (г) (рис. 29), натянутую на контур, к которому предварительно привязана нить так, как показано на рис. 33(а). Пусть эта нить свободно лежит на пленке, т. е. никакие силы на нее не действуют. Будем слегка деформировать контур. Раздвигая правые окружности, мы обнаруживаем, что часть пленки прогибается влево и нить натягивается, превращаясь в сингулярное ребро пленки, сдерживая ее в положении равновесия и не давая деформироваться дальше. Аналогичные события происходят, если раздвигать левые окружности (здесь нить прогнется вправо и натянется). В промежуточном положении, т. е. на пленке (г), нить не натянута. Если в этом критическом положении пленки отцепить нить от одного конца (мы могли не прикреплять нить вторым ее концом к контуру, а держать второй конец нити в руке), придав ей небольшой начальный импульс, то пленка перескочит в состояние (б). Чтобы показать, что при каждом состоянии контура (при небольших u и v) существует неустойчивая пленка типа (в), поступим так. Возьмем пленку (б), прикрепим к одной точке контура нить (рис. 33(б)) и начнем тянуть нить (рис. 33(в)). Скоро мы получим критическое положение пленки. Если продолжить процесс, то произойдет скачок пленки, и нить прогнется влево. Если же, находясь в критическом положении, отпустить нить, придав ей небольшой импульс, то пленка вернется в прежнее положение (б). Рассмотрим теперь состояние нашей системы ($u_0, v_0, (б)$), где u_0 и v_0 невелики, и будем плавно увеличивать v . Пленка (б) сильнее прогибается к центру, и, наконец, наступает такое состояние, когда при дальнейшем малом увеличении v пленка начинает двигаться самостоятельно под действием сил поверхностного натяжения, скачком переходя в состояние (в) при почти не изменившемся состоянии контура.

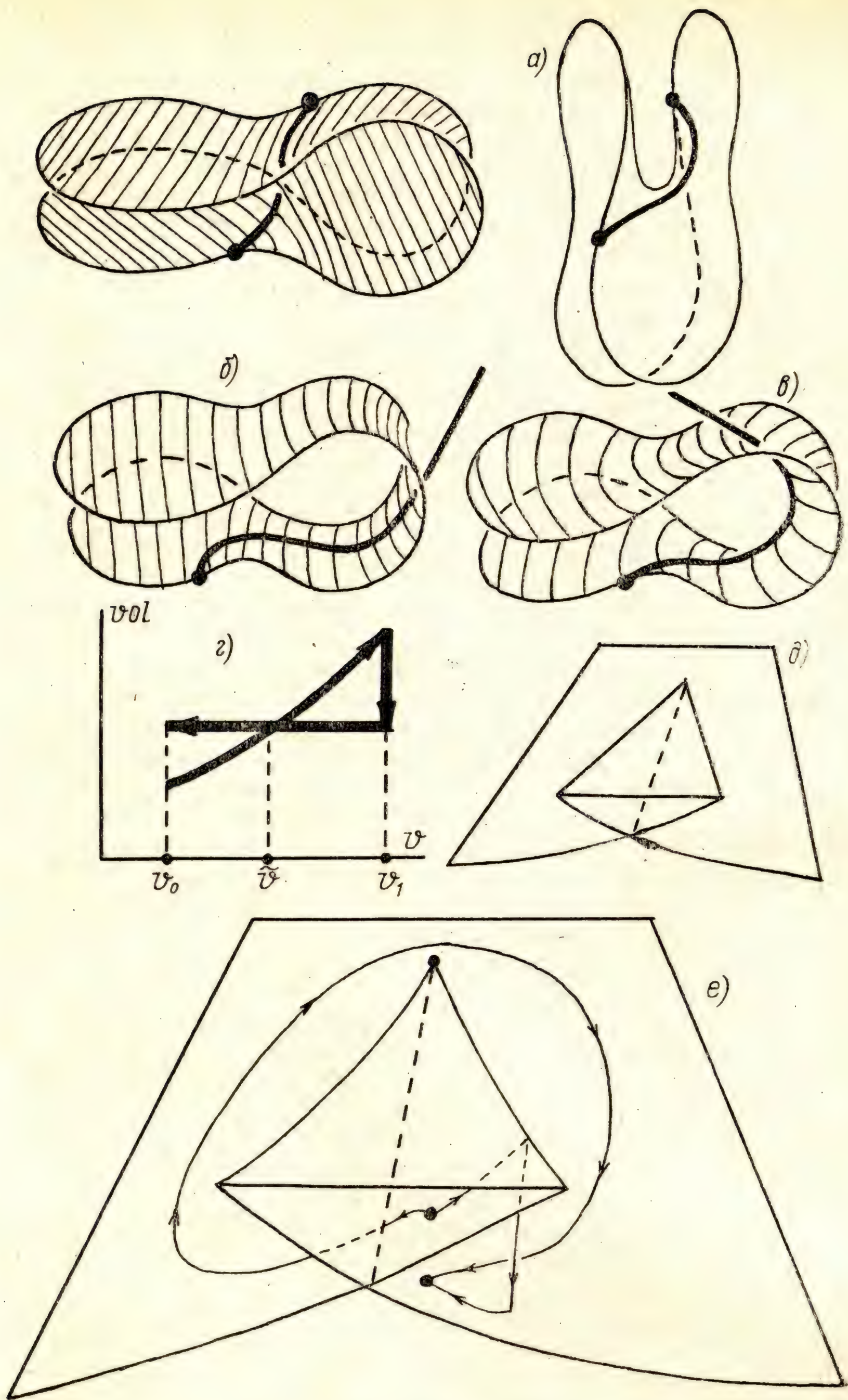


Рис. 33

Если теперь мы будем уменьшать u до начального положения, то с пленкой никаких скачков не происходит. Она сохранит свой тип (в), получившийся в результате скачка. Ввиду недостатка места мы опустим детали и приведем лишь окончательный график изменения площади пленки при указанных деформациях (рис. 33(г)). Значение $\tilde{u}$ соответствует состоянию контура, в котором площадь пленки (б) равна площади пленки (в).

Результат нашего анализа показан на рис. 33(д). Изображен график многозначной функции, описывающей площадь минимальных поверхностей указанного типа. Получившаяся поверхность (и соответствующая ей особенность) называется «ласточкин хвост». Известно, что всякое гладкое отображение трехмерного многообразия в трехмерное многообразие аппроксимируется отображениями, имеющими только особенности Уитни следующего вида: складка, сборка и ласточкин хвост. Ласточкин хвост можно представлять, в частности, как поверхность в трехмерном пространстве многочленов вида $x^4 + ax^2 + bx + c$, состоящую из точек (a, b, c) , отвечающих многочленам с кратными корнями. Вернемся теперь к графику площади мыльных пленок. На рис. 33(е) легко проследить интересную возможность перехода из состояния $(u, v, (б))$ в состояние $(u, v, (в))$ как с помощью скачка, так и непрерывным (плавным) образом. Первый путь — это подъем на сингулярное ребро (путь направо на рис. 33(е)), а затем «падение» из точки на ребре в точку типа (в). Второй путь: плавно спускаемся из точки (б) по листу графика, обходим точку особенности и приходим в точку (в). Итак, неустойчивые пленки типа (г) заполняют (с точки зрения графика их площади) верхний треугольник ласточкина хвоста, а спускающиеся вниз крылья соответствуют устойчивым пленкам (б) и (в).

11. Двумерные минимальные поверхности в трехмерном пространстве и первый принцип Плато.

В 30-е и 40-е годы нашего столетия был достигнут большой прогресс в изучении свойств двумерных минимальных поверхностей в трехмерном пространстве. Замечательные результаты были получены Дугласом, Радом, Курантом и др. Существенную роль в этих исследованиях сыграл принцип Дирихле, который мы сформулируем в следующей форме.

Пусть G — плоская область, ограниченная жордановой кривой γ , и пусть g — функция, непрерывная на замыкании области G , кусочно-гладкая в области G и такая, что интеграл Дирихле $D[g]$ принимает конечное значение на этой функции. Рассмотрим класс всех функций φ , непрерывных на замыкании области G , кусочно-гладких в области G и принимающих на границе области те же значения, что и функция g . Тогда задача нахождения функции φ , на которой интеграл Дирихле $D[\varphi]$ принимает наименьшее значение, имеет единственное решение. Эта же функция является также единственным решением гра-

ничной задачи для уравнения Лапласа $\Delta\varphi=0$ с заранее предписанными граничными значениями $\bar{g}=g|_{\gamma}$ на кривой γ . Напомним, что интегралом (функционалом) Дирихле называется выражение вида $D[\varphi] = \frac{1}{2} \iint (E + G) dudv$, где E, G — коэффициенты первой квадратичной формы (римановой метрики), индуцированной на графике функции φ объемлющей евклидовой метрикой в $\mathbb{R}^3$. Если $r=r(u, v)$ — радиус-вектор поверхности, то $E=\langle r_u, r_u \rangle$, $G=\langle r_v, r_v \rangle$.

В первой половине нашего столетия произошло некоторое смещение акцентов в подходе к решению задачи Плато. Сам Плато сформулировал несколько принципов, которые мы перечислим ниже; здесь же приведем только первый из них.

1-й принцип Плато. Пусть задана поверхность нулевой средней кривизны. Поверхность может быть описана с помощью уравнения, в виде радиус-вектора или при помощи какого-либо геометрического правила, как, например, геликоид. Рассмотрим на ней произвольный кусочно-гладкий замкнутый несамопересекающийся контур, т. е. нарисуем на поверхности замкнутую кривую. Будем считать, что контур ограничивает часть поверхности, причем эта часть устойчива. Изготовим железный проволоочный контур, в точности воспроизводящий нарисованную кривую, затем слегка окислим его разбавленной азотной кислотой и полностью погрузим в мыльную воду (с глицерином). Если мы извлечем контур из жидкости, то среди мыльных пленок, которые могут появиться на этом контуре, обязательно найдется пленка, совпадающая с частью поверхности, ограниченной первоначально нарисованной кривой.

Как видно, Плато исходил из уже заданной минимальной поверхности и лишь затем рассматривал различные контуры, позволявшие реализовывать в виде мыльных пленок те или иные устойчивые части поверхности. С развитием вариационного исчисления доминирующей стала точка зрения, согласно которой первоначальным объектом является контур, и задача состоит в изучении различных минимальных поверхностей, которые можно натянуть на него. Первым по важности вопросом является существование минимальной пленки, затягивающей данный контур. Поскольку изучение произвольных контуров достаточно нетривиально, то сначала был выделен довольно специальный их класс. Рассмотрим поверхности, задающиеся в $\mathbb{R}^3$ гладким графиком однозначной функции $f(x, y)=z$, т. е. допускающие гладкую ортогональную взаимно-однозначную проекцию на двумерную плоскость или на область в ней. Пусть поверхность M проектируется в область G с выпуклой границей γ , т. е. граничная кривая (предполагаемая кусочно-гладкой) выпукла. Если поверхность задается графиком, функционал площади имеет вид $A[f] = \iint_G \sqrt{1 + f_x^2 + f_y^2} dx dy$. Уравне-

нием Эйлера для функционала площади, т. е. уравнением, описывающим критические, стационарные поверхности, является: $(1 + f_x^2) f_{yy} - 2f_x f_y f_{xy} + (1 + f_y^2) f_{xx} = 0$. Это уравнение нелинейное и достаточно сложное, что соответствует прихотливому поведению его решений — минимальных поверхностей — в зависимости от характера граничного контура. Оказывается, для любого замкнутого кусочно-гладкого контура γ' , взаимно-однозначно проектирующегося на плоскость в выпуклую замкнутую кривую γ , существует минимальная поверхность с данной границей, проектирующаяся, следовательно, на область, ограниченную на плоскости кривой γ .

Существование и единственность этого решения в указанном классе поверхностей, допускающих взаимно-однозначную проекцию, есть следствие выпуклости проекции границы. Если от выпуклости области G отказаться, утверждение теоремы становится неверным. Другими словами, существуют гладкие контуры γ' , взаимно-однозначно проектирующиеся на невыпуклую кривую γ в плоскости, и такие, что не существует минимальной поверхности с границей γ' , взаимно-однозначно проектирующейся на область G , ограниченную кривой γ . Простейший пример см. на рис. 34. Минимальная пленка, натягиваемая этот контур γ' , существует, но она, во-первых, не проектируется внутрь кривой γ , а во-вторых, не может быть задана как гладкий график однозначной функции, определенной на плоскости. Ясно, что эти эффекты являются следствием невыпуклости кривой γ .

12. Функционал площади, функционал Дирихле, гармонические отображения и конформные координаты.

Дальнейший существенный прогресс был достигнут Дугласом. Он рассматривал гладкие отображения стандартного двумерного диска $D \subset \mathbb{R}^2$ в $\mathbb{R}^3$. Отнесем диск к евклидовым координатам u, v , тогда его отображение задается радиус-вектором $r = r(u, v)$. Если E, F, G — коэффициенты первой формы поверхности, являющейся образом диска в $\mathbb{R}^3$, то для функционалов Дирихле и площади получаем следующие явные выражения [1, 5]:

$$D[r] = \frac{1}{2} \iint_{D^2} (E + G) du dv, \quad A[r] = \iint_{D^2} \sqrt{EG - F^2} du dv.$$

Экстремальные, критические радиус-векторы для этих функционалов, т. е. решения уравнений Эйлера, устроены так. Экстремальными векторами $r(u, v)$ для функционала площади являются те и только те векторы, для которых средняя кривизна поверхности равна нулю, т. е. минимальные поверхности. Экстремальными векторами для функционала Дирихле являются те и только те векторы, которые являются гармоническими относительно u, v , т. е. $\Delta r = \frac{\partial^2 r}{\partial u^2} + \frac{\partial^2 r}{\partial v^2} = 0$ [1, 5]. Если на

минимальной поверхности задать конформные координаты, то ее радиус-вектор станет гармоническим. Итак, каждая минимальная пленка является гармонической относительно конформных координат. Обратное неверно: поверхность, заметаемая гармоническим радиус-вектором, не обязана быть минимальной. Простейший пример — график вещественной $a(x, y)$ (или мнимой $b(x, y)$) части нелинейной комплексно-аналитической функции $a(x, y) + ib(x, y)$, заданной на плоскости. На

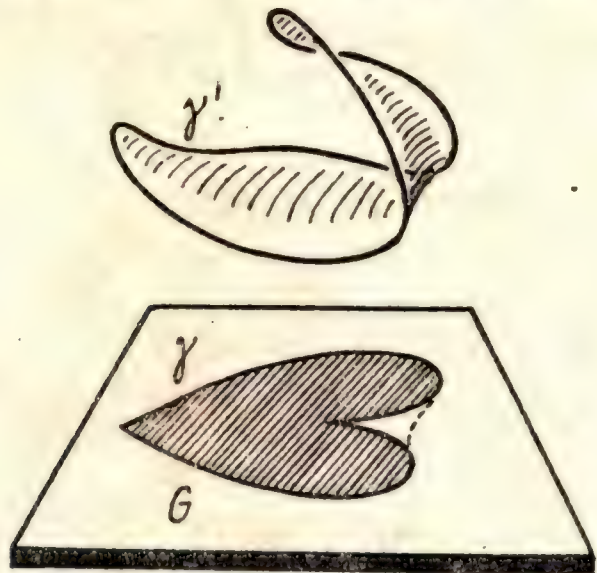


Рис. 34



Рис. 35

рис. 35 условно изображена схема взаимодействия минимальных и гармонических радиус-векторов. Если какой-то радиус-вектор минимален (т. е. соответствующая ему поверхность минимальна), то и все другие радиус-векторы, полученные из исходного регулярной заменой переменных, также минимальны. При таких заменах поверхность не меняется, а средняя кривизна — инвариант замены координат. В каждом таком классе эквивалентных друг другу (относительно замен переменных) минимальных радиус-векторов обязательно имеется гармонический. На рис. 35 показано, что каждый класс эквивалентности пересекается с множеством гармонических векторов.

Функционалы площади и Дирихле связаны неравенством: $D[r] \geq A[r]$, причем равенство достигается в том и только в том случае, когда $E=G$, $F=0$, т. е. когда координаты u, v конформны. Это вытекает из очевидного неравенства

$$\frac{E+G}{2} \geq \sqrt{EG-F^2}.$$

Некоторое аналитическое преимущество функционала Дирихле перед функционалом площади заключается в том, что его подынтегральное выражение не содержит радикалов. Аналогичная картина наблюдается и в теории одномерных функционалов длины пути и действия пути [2]. Здесь экстремалами функционала действия $\int \langle \dot{\gamma}, \dot{\gamma} \rangle dt$ являются геодезические, а экстремалами функционала длины $\int \sqrt{\langle \dot{\gamma}, \dot{\gamma} \rangle} dt$ — всевозможные регулярные параметризации геодезических. В одно-

мерном случае у функционала длины дуги «больше» экстремалей, чем у функционала действия. В двумерном случае картина сложнее. Наиболее важным результатом, доказанным Дугласом, является теорема существования минимальной поверхности в классе поверхностей фиксированного топологического типа. Прежде чем сформулировать теорему, напомним, что каждое гладкое компактное замкнутое связное двумерное многообразие диффеоморфно либо сфере с некоторым количеством ручек, либо сфере, заклеенной некоторым количеством пленок Мебиуса (рис. 36). Ручкой называется обыкновенный цилиндр, приклеиваемый своими двумя краями к двум дыркам в двумерной сфере. Пленкой Мебиуса называют обычный лист Мебиуса, приклеенный по своей границе (по окружности) к границе дырки в сфере. Иногда для наглядности пленку Мебиуса реализуют в виде скрещенного колпака (рис. 36). Для этого нужно так изогнуть границу листа Мебиуса, чтобы она превратилась в стандартную окружность. За это придется «заплатить» усложнением конфигурации листа Мебиуса и появлением на нем самопересечений. Теперь мы можем изложить результат Дугласа. Предположим, что для двумерного многообразия M данного топологического типа (т. е. с заданным числом ручек или пленок Мебиуса) с краем существует гладкое отображение $f: M \rightarrow \mathbb{R}^3$ с конечной площадью, т. е. $A[f] < \infty$, причем край поверхности отображается на заданный набор замкнутых жордановых кривых в $\mathbb{R}^3$. Предположим, что минимум (точнее, точная нижняя грань) площадей таких погружений поверхности в $\mathbb{R}^3$ с фиксированным краем строго меньше, чем минимум площадей всех погружений в $\mathbb{R}^3$ (с тем же краем) всех поверхностей, получающихся из исходной выбрасыванием одной ручки или выбрасыванием одной пленки Мебиуса. Тогда существует гладкое отображение $f_0: M \rightarrow \mathbb{R}^3$, являющееся минимальным, т. е. оно реализует минимум площади в классе всех погружений поверхности в $\mathbb{R}^3$ с заданной границей. Эта поверхность имеет нулевую среднюю кривизну во всех своих неособых точках. Прокомментируем условие существования хотя бы одного погружения M в $\mathbb{R}^3$ с конечной площадью. Аналогичное условие накладывалось и при формулировке принципа Дирихле. Дело в том, что существуют кривые, на которые нельзя натянуть ни одной поверхности с конечной площадью (рис. 37). На рис. 37 изображены две ортогональные проекции пространственной кривой на две плоскости. Уравнение кривой в стан-

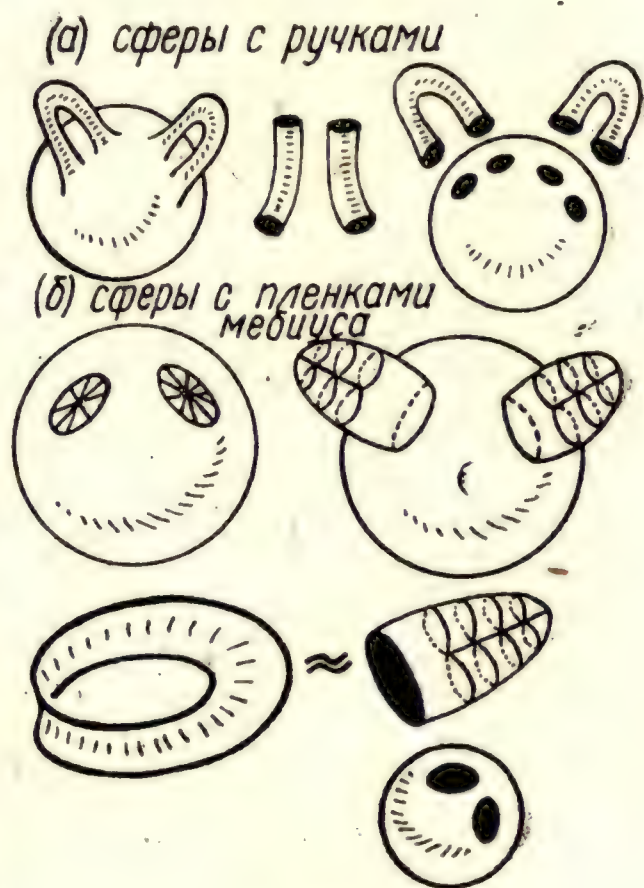


Рис. 36

дартных сферических координатах ρ , θ , φ записывается так: $\rho = \cos \theta$, $\varphi = \operatorname{tg}^5 \theta$. Здесь ρ — длина радиус-вектора; φ — полярный угол на плоскости x, y ; а $\frac{\pi}{2} - \theta$ — угол, образуемый радиус-вектором с вертикальной осью z . Когда θ стремится к $\pm \pi/2$, то $|\varphi| \rightarrow \infty$, $\rho \rightarrow 0$, что и дает картину двух проекций. Прямой подсчет показывает, что эта кривая не является границей никакой погруженной в $\mathbb{R}^3$ поверхности с конечной площадью. Более того, существуют примеры таких экзотических кривых, каждая точка которых имеет похожую сложную структуру, и, следовательно, заклейка любого, сколь угодно малого участка границы требует построения поверхности бесконечной площади. Остановимся на свойствах минимальных поверхностей, существование которых утверждается в теореме Дугласа.

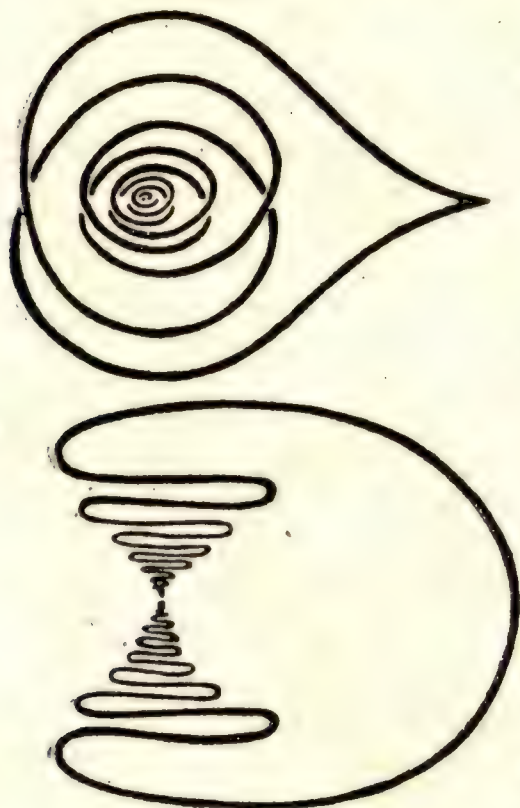


Рис. 37

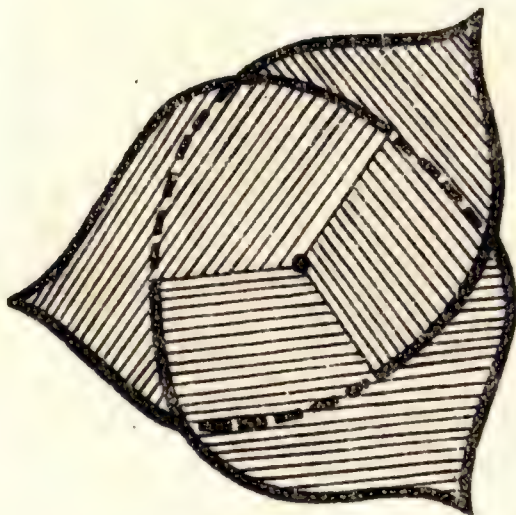
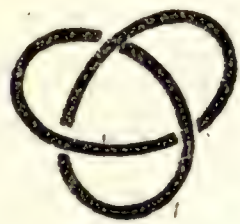


Рис. 38

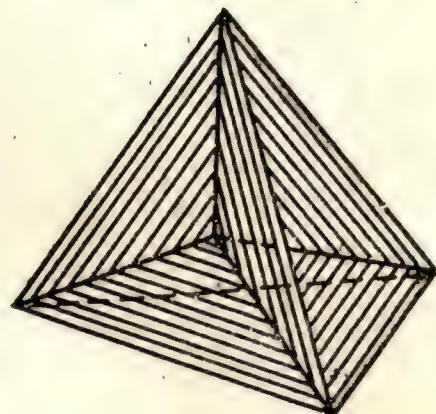
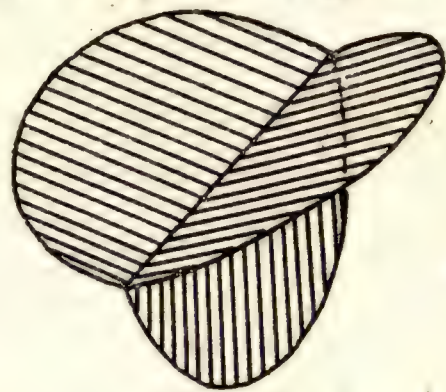


Рис. 39

Поскольку рассматриваются гладкие отображения поверхности в $\mathbb{R}^3$, являющиеся погружениями во всех точках, кроме, быть может, множества точек меры нуль, то минимальные пленки могут иметь самопересечения и даже точки ветвления. Ясно, что существуют контуры, для которых минимальной пленки в классе вложений не существует. Например, простейший заузленный контур (трилистник) см. на рис. 38. Он не затягивается в $\mathbb{R}^3$ никаким вложенным диском. В то же время существует гладкое отображение диска D^2 в $\mathbb{R}^3$ (с самопересечениями, реализующее минимум площади в классе отображений диска с данной границей. Это утверждение для отображений диска верно, кстати, для любого жорданова замкнутого контура, гомеоморфного окружности. На рис. 38 показан также пример минимальной пленки, затягивающей вложенную в $\mathbb{R}^3$ окружность, заузленную в виде трилистника. Эта поверхность имеет не только самопересечения (три отрезка, сходящиеся в одной точке), но и точку ветвления (в центре). Однако, эта поверхность неустойчива.

13. Особые, сингулярные точки минимальных поверхностей и три принципа Плато.

Важным свойством минимальных поверхностей общего типа является частое наличие на них особых, сингулярных точек, т. е. таких, около которых поверхность устроена более сложно, чем обычный двумерный плоский диск малого радиуса. Наличие таких особенностей является «типичной ситуацией» в том смысле, что в случае произвольного граничного контура затягивающая его мыльная пленка «почти наверняка» будет иметь сингулярности. Множества особых точек могут быть достаточ-

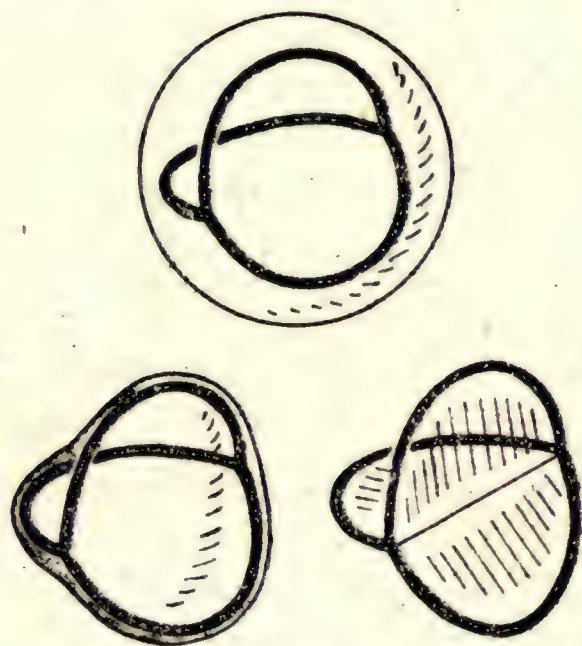


Рис. 40

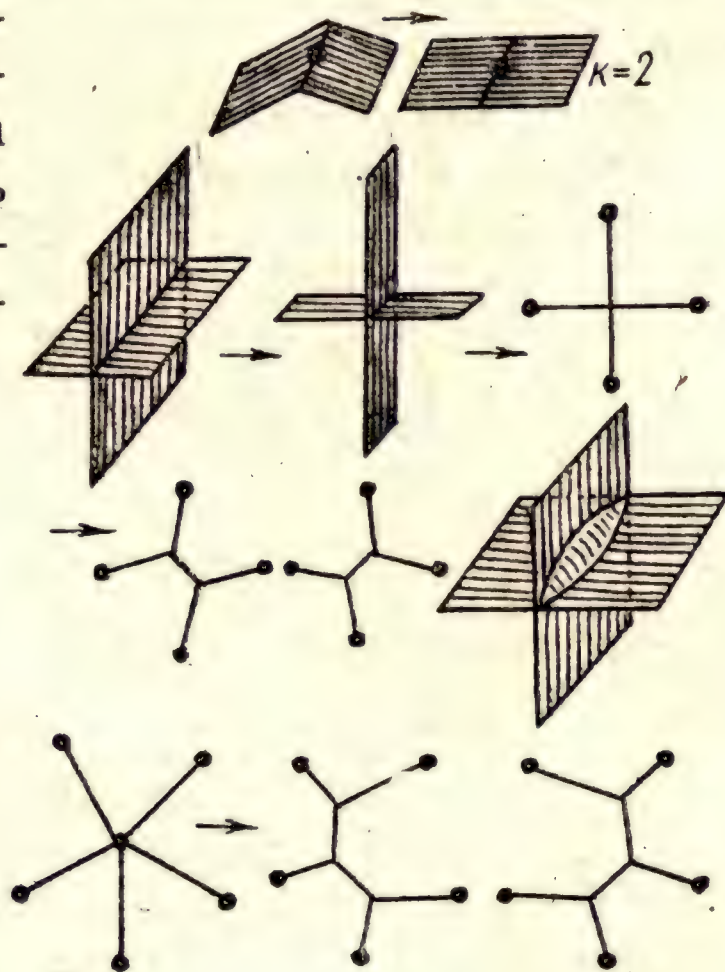


Рис. 41

но сложными и интересными. Довольно часто особые точки не являются изолированными, а заполняют целые отрезки, к которым подходит несколько листов пленки. Каков характер особых точек? Еще Плато обнаружил экспериментальным путем три обстоятельства, которые мы условно назовем 2-м, 3-м и 4-м принципами Плато.

2-й принцип Плато. Минимальные поверхности (нулевой средней кривизны) и системы мыльных пузырей (постоянной положительной средней кривизны) состоят из фрагментов гладких поверхностей, сходящихся, стыкующихся друг с другом вдоль гладких дуг.

3-й принцип Плато. Куски гладких поверхностей, из которых состоят пленки постоянной средней кривизны, могут стыковаться только следующими двумя способами: а) три гладких листа сходятся на одной гладкой кривой, б) шесть гладких листов (вместе с четырьмя особыми кривыми) сходятся в одной точке, вершине.

4-й принцип Плато. В том случае, когда три листа поверхности стыкуются по общей дуге, ребру, они образуют друг с другом равные углы в 120° . В том случае, когда четыре особые кривые (и шесть листов) сходятся в одной вершине,

эти ребра образуют друг с другом равные углы приблизительно в 109° .

Соответствующие геометрические конструкции показаны на рис. 39. Каждая из изображенных мыльных пленок реализуется в виде поверхности, затягивающей соответствующий контур, так как пленки устойчивы. Полезно представлять себе один из механизмов образования трех листов поверхности, сходящихся на общем ребре под углами в 120° . Рассмотрим контур, заключенный внутри большого мыльного пузыря (рис. 40). Постепенно выпуская воздух из пузыря, мы видим, что он начинает обволакивать контур и, наконец, схлопывается на поверхность требуемой формы. Будем говорить, что точка минимальной поверхности имеет кратность k , если в этой точке сходится k листов поверхности. Регулярная точка является точкой кратности два: два листа сходятся по общему ребру под углом в 180° , т. е. поверхность является гладкой в окрестности данной точки. Из рис. 41 видно, что два листа минимальной поверхности не могут сходиться на общем ребре под углом, отличным от π , в противном случае есть деформация пленки, уменьшающая ее площадь. Если мы допускаем к рассмотрению поверхности, имеющие особые точки, то с математической точки зрения существуют пленки, средняя кривизна которых почти всюду равна нулю (в неособых точках), а кратность особых точек может быть больше чем три. Например, две ортогональные плоскости, пересекающиеся по прямой, реализуют такую поверхность. Кратность каждой точки самопересечения равна здесь 4 (рис. 41). Тем не менее попытки построить реальную мыльную пленку, у которой на каком-либо сингулярном отрезке сходилась бы число листов, большее чем три, терпят неудачу. Оказывается, любая поверхность нулевой кривизны с самопересечениями является неустойчивой, т. е. не может быть, в частности, реализована устойчивой мыльной пленкой. Это утверждение носит даже не глобальный, а локальный характер: неустойчивой является уже сколь угодно малая окрестность каждой точки самопересечения. В самом деле, предположим, что на каком-то отрезке самопересечения поверхности сошлись 4 листа. Можно считать, что они попарно ортогональны, так как в противном случае неустойчивость пленки очевидна. Вырежем часть пленки, расположенную около отрезка самопересечения, используя 1-й принцип Плато, и фиксируем границу получившихся двух тонких полосок пленки. Рассмотрим сечение пленки плоскостью, ортогональной сингулярному отрезку. Четырехкратная особая точка распадается в объединение двух трехкратных точек, поскольку длина кривой, состоящей первоначально из двух ортогональных отрезков, больше, чем длина каждой из кривых, показанных на рис. 41 (проверьте!). Минимизируя длину кривой, являющейся плоским сечением, мы минимизируем площадь выделенного куса поверхности. На рис. 41 этот процесс показан в $\mathbb{R}^3$ в непосредственной близости от се-

чения. Итак, мы построили деформацию пленки, уменьшающую ее площадь. Так как пленка находится теперь в устойчивом равновесии, в частности, малые ее возмущения лишь увеличивают площадь, то равнодействующая всех сил, действующих на каждую точку пленки, равна нулю. Это относится и к особым точкам. В положении равновесия в каждой такой точке сходится ровно три листа, поэтому ясно, что они сходятся под равными углами $2\pi/3$. Аналогичные рассуждения применимы и тогда, когда на общем ребре сходится больше четырех листов. Поверхность снова перестраивается, занимая положение с меньшей площадью, в результате чего сингулярное ребро высокой кратности распадается в несколько трехкратных ребер (рис. 41). Кстати, отметим еще раз, что на один контур можно натянуть, вообще говоря, несколько мыльных пленок. На рис. 41 показано, что особые точки большой кратности могут несколькими способами распадаться в объединение трехкратных точек.

Выше мы рассматривали сечение поверхности плоскостью, ортогональной сингулярному ребру. Построенная деформация была равномерной во всех точках ребра, близких к сечению, т. е. эта деформация смещает границу ортогональных полосок пленки (рис. 41). Однако легко построить деформацию, уменьшающую площадь пленки и при условии жесткой фиксации границы части пленки, показанной на рис. 41. Четырехкратное ребро распадается в объединение двух трехкратных изогнутых ребер, концы которых неподвижны. Итак, мы построили монотонное сокращение пленки, уменьшающее ее площадь при неподвижной границе. Вклеивая этот кусок пленки в исходную пленку большого размера, мы и получаем искомую сокращающую деформацию. Этот же процесс можно продемонстрировать на примере гладкого погружения двумерного диска в $\mathbb{R}^3$ (рис. 42). Узкая полоска диска, поднявшись вверх, ортогонально пересекает основную плоскость пленки. Попытка реализовать эту поверхность в виде реальной мыльной пленки сразу обнаруживает ее неустойчивость. Четырехкратное ребро самопересечения распадается в сумму двух изогнутых трехкратных ребер с общими концами.

Эти задачи связаны с двумерной задачей Штейнера (Steiner): как найти нить наименьшей длины, соединяющей конечное множество точек на плоскости? При этом нить может разветвляться, иметь точки кратности три и выше. С помощью мыльных пленок эта задача решается экспериментально. Нужно взять две близкие плоскости, изготовленные из тонкого органического стекла, соединить их вертикальными столбиками-отрезками, расположенными в тех точках плоскости, через которые должна пройти искомая нить. Опуская конструкцию в мыльную воду и извлекая обратно, мы получаем пленку, натянутую на столбики и реализующую минимум длины ее плоского сечения. При этом нить содержит точки кратности не большей, чем три.

Аналогичные механизмы действуют и в тех случаях, когда в формировании мыльной пленки играет роль давление воздуха, например когда в ней образуются пузыри, ограничивающие замкнутый объем, заполненный воздухом. В этом случае пленка занимает положение, компенсирующее давление воздуха, заключенного внутри. В общем случае, комбинируя затягивание контура мыльной пленкой с процессом выдувания пузырей, можно получать реальные пленки, заключающие в себе как

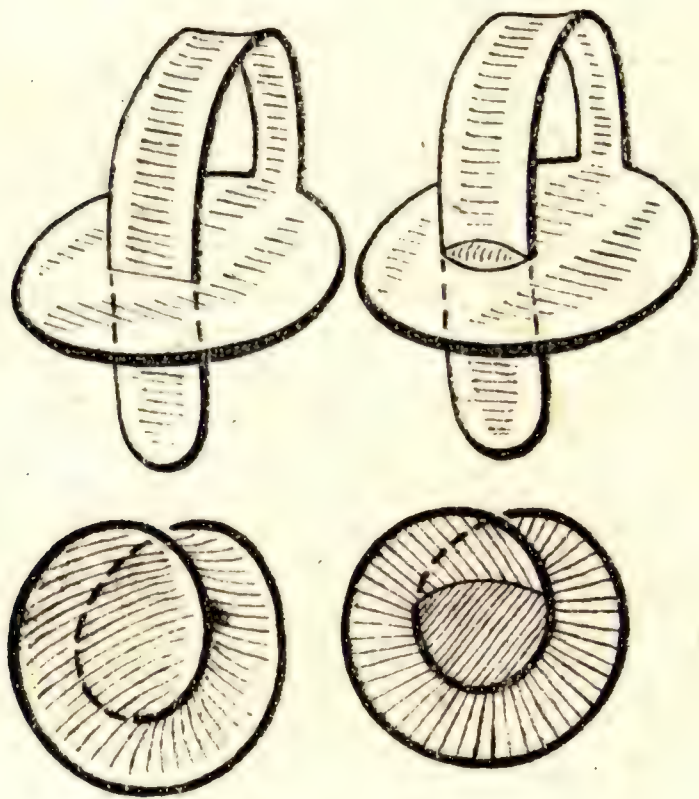


Рис. 42

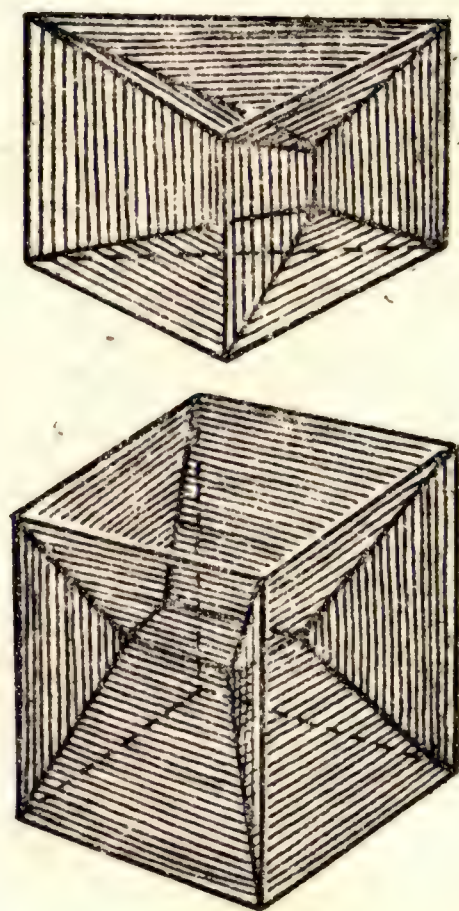


Рис. 43

пузыри, так и поверхности нулевой средней кривизны. Эти эффекты наблюдаются, например, при взбивании мыльной пены.

На рис. 43 показаны достаточно сложные граничные контуры и некоторые из минимальных пленок, натянутые на них [11]. Видно, что взаимодействие различных листов поверхности часто носит прихотливый характер и усложняется по мере усложнения граничного контура. Иногда мыльные пленки, включающие в себя пузыри, ограничивающие замкнутые объемы воздуха, возникают самопроизвольно при извлечении из жидкости достаточно сложных контуров.

14. Реализация минимальных поверхностей в живой природе.

Еще начиная с XVIII столетия было известно, что многие конкретные задачи физики, химии, биологии сводятся к анализу минимальных поверхностей. Приведем пример из биологии. Оказывается, минимальные пленки широко распространены в природе как наиболее экономичные поверхности, формирующие скелеты некоторых живых организмов. Наиболее эффективным примером являются скелеты радиолярий — микроскопических морских животных, имеющих самые разнообразные и экзотические формы.

По-видимому, первым, кто обратил особое внимание на то, что поверхностное натяжение играет существенную роль в создании форм этих живых существ, был Томпсон (D'Arcy Wentworth Thompson), автор книги «Рост и форма». Радиолярии состоят из небольших комочков протоплазмы, заключенных в пенообразные формы, наподобие мыльных пузырей и пленок. Поскольку эти организмы достаточно сложны, то минимальные поверхности, входящие в их состав, имеют, вообще говоря, много точек ветвления и трехкратных особых ребер, на которых и концентрируется основная масса жидкости, входящей в состав организма. Эта концентрация жидкости вдоль особых ребер и около особых точек хорошо видна на реальных мыльных пленках. Жидкость стекает с пленки свободно до тех пор, пока не встречает ребро, на котором сходятся листы пленки. Здесь жидкость тормозится и оседает, образуя водяные отрезки, визуально выявляющие особые ребра минимальной поверхности. Этот же процесс фактически происходит и у радиолярий. Концентрация жидкости вдоль ребер ветвления приводит к тому, что твердые фракции морской воды и соли оседают вдоль этих ребер и постепенно образуют твердый скелет животного. Наглядно геометрию этого скелета можно представить, если в массе мыльной пены обратить внимание на сингулярные ребра, по



Рис. 44(а)

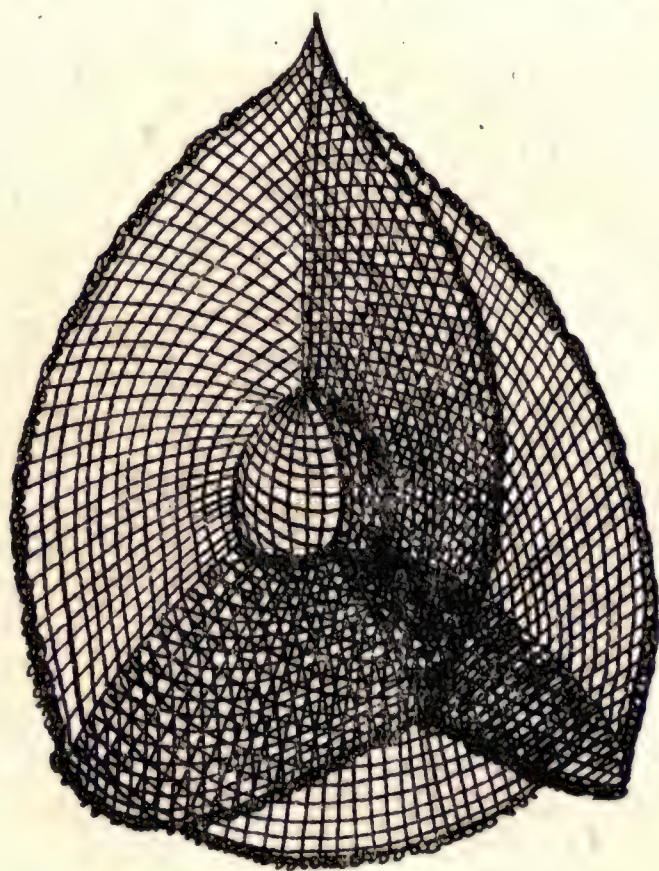


Рис. 44(б)

которым стыкуются разные пузыри. Эти отрезки соединяются друг с другом иногда весьма причудливым образом и образуют «жидкостный костяк» пены. После гибели животного мягкие ткани постепенно исчезают, разрушаются и остается твердый скелет, образовавшийся описанным образом. Он представляет собой не что иное, как наглядное застывшее изображение (фотографию) системы ребер ветвления и особых точек, возникших в сложной минимальной поверхности, включавшей

в себя и замкнутые объемы. Кроме того, в формировании твердого скелета радиолярии определенную роль играют и стенки клеток, также аккумулирующие в себе морские соли. На рис. 44 показаны три скелета радиолярий. Рисунок взят из книги:

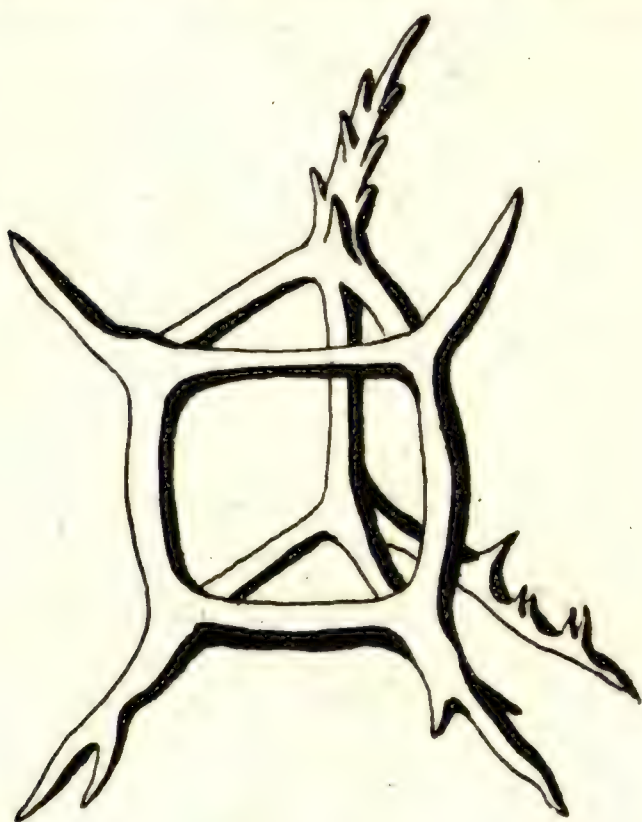


Рис. 44(в)

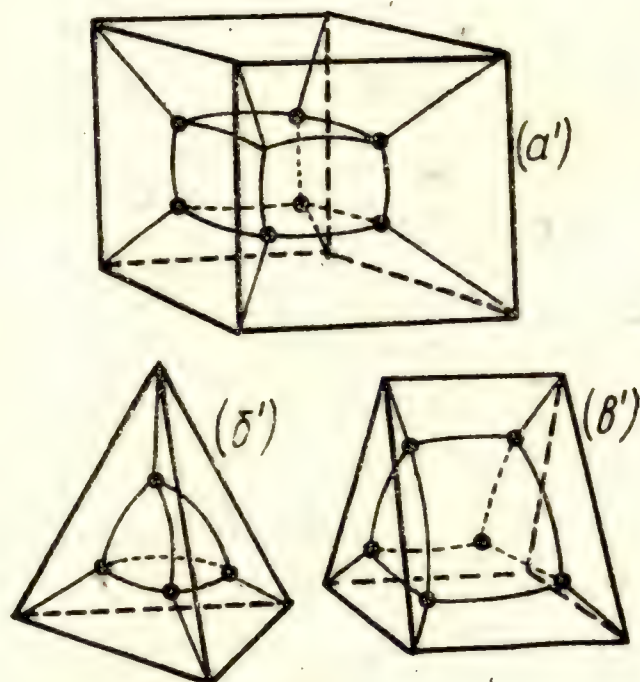


Рис. 44(а)'—(в)'

Ernst Haeckel "Report on the Scientific Results of the Voyage of the HMS Challenger during the Years 1873—1876". Здесь же изображены три минимальные поверхности с включением мыль-

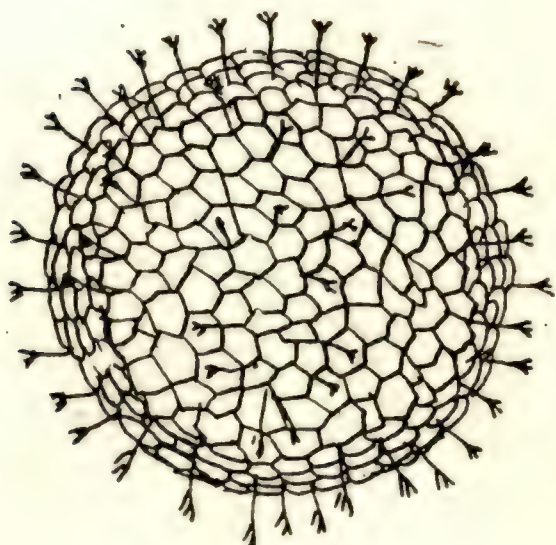


Рис. 45(а)

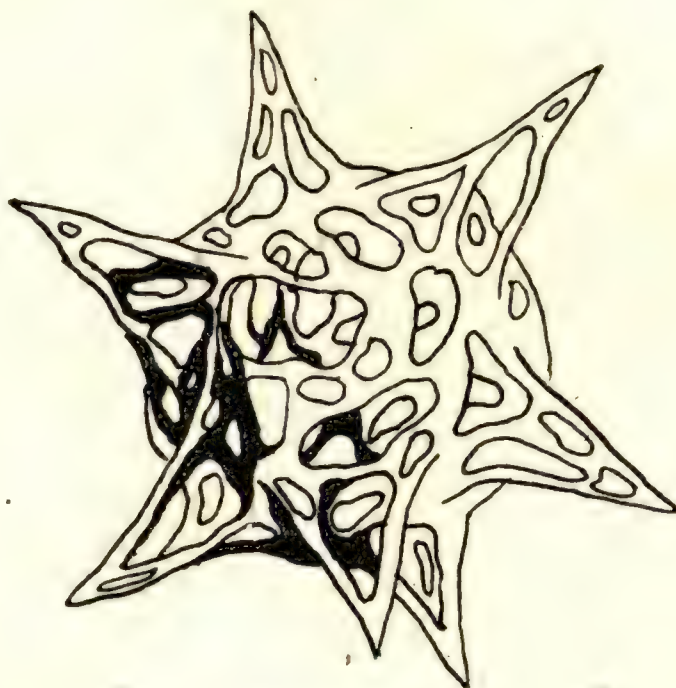


Рис. 45(б)

ных пузырей, натянутых на контуры, являющиеся системой ребер обычного тетраэдра, куба и призмы. Сходство форм реальных организмов с указанными пленками поразительно. Видно, что скелеты радиолярий довольно точно воспроизводят картину ребер ветвления минимальных поверхностей. Еще два экзотических скелета радиолярий см. на рис. 45.

И это далеко не единственный пример той роли, которую минимальные поверхности играют в окружающей нас живой природе. Важна роль таких пленок и в химии, где поверхность-

ные взаимодействия на границах раздела различных сред определяют характер и скорость многих химических превращений. Примерами минимальных поверхностей могут служить всем нам хорошо известные мембраны, встречающиеся как в различных природных образованиях, так и в некоторых областях человеческой деятельности. Это и барабанная перепонка в ухе, это мембраны, служащие границами живых клеток и т. п. Основное свойство мембраны — стремление занять положение, оптимальное с энергетической точки зрения для данных внешних условий. Поэтому форма мембраны в значительной степени определяется внешними взаимодействиями и, напротив, по форме минимальной поверхности можно судить о том, какие воздействия являются оптимальными для организма. Реальные мембраны обладают еще целым рядом важных свойств, например избирательной проницаемостью или полупроницаемостью, позволяющей клеткам производить обмен ионами.

Как мы видим, с минимальными поверхностями связан и целый ряд интересных математических проблем. Их корректная постановка и исследование позволяют лучше понять этот удивительный, хотя и простой с виду природный феномен.

§ 2. Топологические свойства минимальных поверхностей

1. О различных подходах к понятиям поверхности и границы поверхности.

В этом параграфе мы попытаемся вкратце описать развитие исследований по задаче Плато начиная с 60-х годов нашего столетия, когда в этой области произошел существенный качественный скачок, вновь привлечший особое внимание специалистов к классической задаче изучения минимальных поверхностей. Мы отнюдь не претендуем даже на хотя бы беглый обзор всех этих новых направлений, связанных с замечательными результатами таких авторов, как: Федерер, Флеминг, Альмгрен, Миранда, Райфенберг, Морри, Бомбьери, Джусти, Де Джорджи, Саймонс, Лоусон, Аллард и многие другие. Достаточно полная информация и библиография содержатся, например, в [7—16, 18, 19], куда мы и отсылаем заинтересованного читателя. Мы ограничимся изложением лишь новых качественных эффектов топологического характера, позволивших значительно расширить наши представления о свойствах минимальных поверхностей.

В предыдущем параграфе мы специально не акцентировали внимание читателя на обсуждении понятий «пленки» и «границы», ограничившись либо интуитивным представлением, либо рассмотрением класса двумерных поверхностей с краем, гладко отображаемых в $\mathbb{R}^3$. Однако понятия поверхности и границы являются не столь простыми и однозначными, как это может показаться на первый взгляд. Если в случае одномерного кон-

тура можно ограничиваться изложенными выше представлениями, то в случае многомерной задачи Плато, связанной с изучением функционала многомерного объема, сама формулировка указанных понятий требует специального исследования. Прогресс, достигнутый начиная с 60-х годов, в значительной мере обусловлен выработкой достаточно гибких понятий поверхности и ее границы.

Из физических свойств двумерных минимальных поверхностей следует, что в них не должно быть «дырок», проколов, не взаимодействующих с границей пленки. В противном случае такая дырка под действием сил поверхностного натяжения бу-

дет расширяться до тех пор, пока вся пленка или ее часть не стянется, схлопнется на граничный контур. Это свойство легко проверяется на реальном опыте. Нужно только быстро проколоть мыльную пленку достаточно толстой проволокой или стержнем. Указанное свойство мыльных пленок и кладется в конечном итоге в основу математического определения границы и свойства пленки затягивать, заклеивать границу. Конечно, соответствующая математическая идеализация является упрощением реальности, поскольку существуют (и устойчивы), физические мыльные пленки, повисающие на проволоочном контуре, однако оставляющие некоторые его участки свободными. Другими словами, мыльная пленка может быть устойчивой и

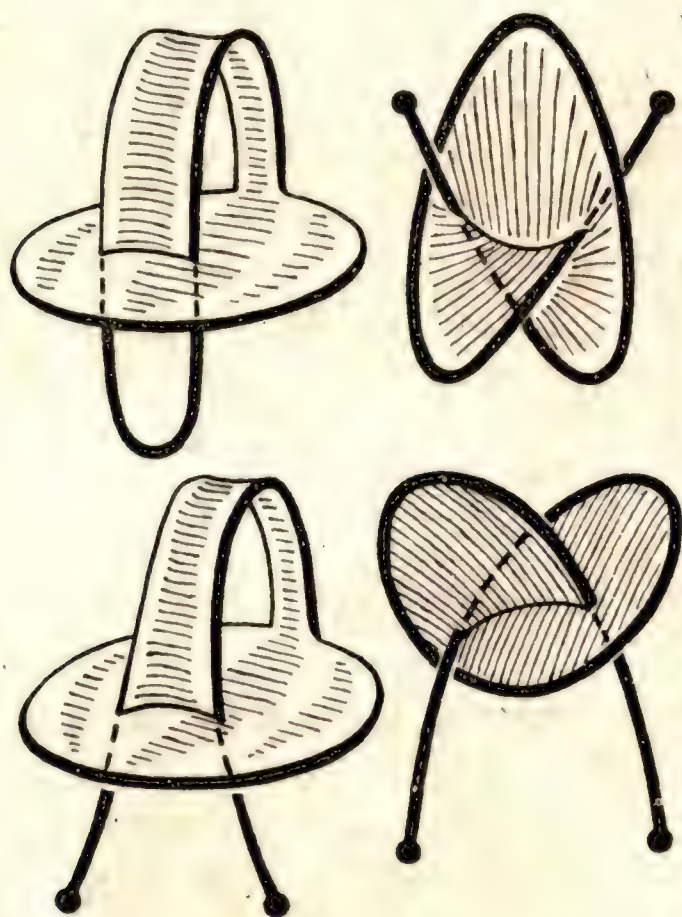


Рис. 46

в то же время крепиться не ко всему контуру, а лишь к его части. Это объясняется тем, что реальная проволока имеет конечную толщину, что стабилизирует в некоторых случаях математически невозможную ситуацию (рис. 46). Далее мы будем рассматривать замкнутые контуры, однако следует иметь в виду (хотя мы и не будем этим заниматься), что существуют реальные мыльные пленки, затягивающие незамкнутые контуры (рис. 46). Этот контур незамкнут, поскольку распрямляется в отрезок. Здесь проволока должна быть достаточно толстой по сравнению с размерами пленки. С математической точки зрения не существует минимальной поверхности (в разумном смысле), затягивающей отрезок, вложенный без самопересечения в $\mathbb{R}^3$.

Исследование критических (минимальных) поверхностей для функционала многомерного риманова объема показывает, что класс гладко вложенных или погруженных подмногообразий, на котором определен функционал, недостаточно широк для того, чтобы содержать решение вариационной задачи. Другими

словами, минимальная поверхность может не принадлежать этому классу. В связи с этим Райфенберг отметил: «Хотя интуитивно ясно, что любое множество, являющееся поверхностью минимальной площади, в некотором смысле должно быть локально хорошо устроено, тем не менее это утверждение следует доказывать, и доказательство это не может быть проведено, если мы с самого начала будем рассматривать лишь локально хорошо устроенные поверхности. Другими словами, было бы полезно исследовать структуру множеств минимальной площади, не предполагая заранее, что они похожи на многообразия».

2. Гомологическая граница поверхности и роль группы коэффициентов.

Поскольку мы хотим расширить понятие поверхности, следует соответственно расширить и понятие ее площади или объема в многомерном случае. В качестве основного функционала этого типа, определенного на замкнутых k -мерных подмножествах в $\mathbb{R}^n$, можно взять k -мерную сферическую меру Хаусдорфа, которую будем обозначать $\text{vol}_k X$. Ее определение и свойства см., например, в [7, 8, 20]. Мы не будем здесь давать формальное определение этой меры, ограничившись утверждением, что если подмножество X является гладким подмногообразием или стратифицированным алгебраическим подмногообразием в $\mathbb{R}^n$, то $\text{vol}_k X$ совпадает с обычным k -мерным римановым объемом подмногообразия. В качестве «поверхностей» можно теперь рассмотреть измеримые, т. е. имеющие конечную меру Хаусдорфа, компактные подмножества в $\mathbb{R}^n$. Оказывается, этот класс уже достаточно широк и содержит в себе минимальные поверхности. Осталось сформулировать понятие границы. Фиксируем какую-либо теорию гомологий с коэффициентами в группе G . Если поверхность X является конечным клеточным комплексом или гладким многообразием, то в качестве группы q -мерных гомологий $H_q(X, G)$ можно брать обычные клеточные или симплициальные гомологии [1, 4]. Если же X является лишь измеримым компактным множеством, то в качестве $H_q(X, G)$ следует рассматривать так называемые спектральные гомологии (гомологии Чеха). В случае конечных клеточных комплексов и гладких многообразий эти группы гомологий совпадают с группами клеточных (симплициальных) гомологий. Спектральные гомологии являются естественным обобщением этих последних групп на случай измеримых компактов [7, 8]. Мы не будем поэтому вникать здесь в технические детали построения спектральных гомологий, поскольку эти подробности нам не потребуются, а заинтересованный читатель всегда может их восстановить по стандартной литературе.

Естественно считать, что поверхность X (в широком смысле) целиком заклеивает границу A , если A содержится в X и поверхность не имеет «дырок», граница которых имеет ту же размерность, что и A . Оказывается, группы гомологий хорошо

приспособлены для того, чтобы математически формализовать это интуитивное представление. Будем говорить, следуя Дж. Ф. Адамсу и Райфенбергу, что замкнутое k -мерное множество $X \subset \mathbb{R}^n$ является G -поверхностью, затягивающей, заклеивающей замкнутое $(k-1)$ -мерное множество A для данной абелевой группы коэффициентов G , если, во-первых, A содержится в X , а во-вторых, гомоморфизм $i_* : H_{k-1}(A, G) \rightarrow H_{k-1}(X, G)$, индуцированный вложением $i : A \rightarrow X$, тождественно равен нулю, т. е. вся группа $H_{k-1}(A, G)$ отображается в ноль.

Другими словами, все $(k-1)$ -мерные циклы, содержащиеся в границе A , должны «раствориться», аннулироваться в поверхности X . В этом смысле A является гомологической границей X .

Если мы интересуемся минимальными поверхностями в $\mathbb{R}^3$, то A считается границей X , если гомоморфизм $i_* : H_1(A, G) \rightarrow H_1(X, G)$ тривиален. Данное выше определение допускает эквивалентную (но иногда полезную) переформулировку. Если рассмотреть фрагмент точной последовательности пары пространств (X, A) , а именно

$$H_2(X, A, G) \xrightarrow{\partial} H_1(A, G) \xrightarrow{i_*} H_1(X, G),$$

то условие тривиальности гомоморфизма i_* эквивалентно условию, что граничный гомоморфизм $\partial : H_2(X, A, G) \rightarrow H_1(A, G)$ является эпиморфизмом. Это определение границы поверхности в значительной степени согласуется с нашим интуитивным представлением о границе. В самом деле, рассмотрим некоторые частные случаи.

1) Пусть M^k — гладкое компактное ориентируемое многообразие с краем A , являющимся гладким $(k-1)$ -мерным ориентированным многообразием. Здесь мы имеем классическую ситуацию: многообразие с краем в смысле дифференциальной топологии. Рассмотрим вложение $i : A \rightarrow M$, тогда гомоморфизм $H_{k-1}(A, G) \rightarrow H_{k-1}(M, G)$ тривиален и A является границей M в гомологическом смысле.

2) Пусть A — произвольный конечный клеточный комплекс и X — конус над A (рис. 47). Тогда индуцированный гомоморфизм $H_{k-1}(A, G) \rightarrow H_{k-1}(X, G)$ тривиален (для всех k). Это вытекает из того, что вся «поверхность» X стягивается по себе в точку.

Понятие G -поверхности с гомологической границей A существенно зависит от выбора группы G . Множество A может быть границей X для одной группы и не быть границей для другой. Приведем примеры.

3) Если A — окружность, а X — лист Мебиуса (неориентируемая поверхность), то индуцированный вложением гомоморфизм $i_* : H_1(A, \mathbb{Z}) \rightarrow H_1(X, \mathbb{Z})$ имеет нулевое ядро, т. е. является мономорфизмом $i_* : \mathbb{Z} \rightarrow \mathbb{Z}$, $i_* : 1 \rightarrow 2$, т. е. i_* умножает каждый цикл на 2. Это следует из того, что с гомотопической точки зрения вложение $i : A \rightarrow X$ эквивалентно отображению

окружности A на себя со степенью два (рис. 47). Итак, лист Мебиуса не является $\mathbb{Z}$ -поверхностью с границей окружность, хотя является гладким многообразием с границей окружность. Однако лист Мебиуса является $\mathbb{Z}_2$ -поверхностью с границей окружность. Для группы $G = \mathbb{Z}_2$ получаем, что гомоморфизм $i_* : H_1(A, \mathbb{Z}_2) \rightarrow H_1(X, \mathbb{Z}_2)$ тривиален, поскольку приведение по модулю два аннулирует образ целочисленной образующей в $H_1(X, \mathbb{Z}_2)$.

4) Пусть A — окружность, а X — тройной лист Мебиуса, получающийся так. Рассмотрим незаузленную окружность в $\mathbb{R}^3$

и «трилистник»: три отрезка одинаковой длины, сходящиеся в некоторой точке окружности ортогонально ей и образующие друг с другом равные углы $2\pi/3$. Будем перемещать эту фигуру вдоль окружности, оставляя трилистник ортогональным ей, но вращая вокруг оси так, чтобы после совершения полного оборота отрезки подверглись перестановке: $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$ (рис. 47).

Границей получившейся поверхности является окружность, три раза обходящая вокруг вертикальной оси. Поверхность не является гладким многообразием, так как содержит целую окружность тройных сингулярных точек. Следовательно, тройной лист Мебиуса образуется по тому же рецепту, что и обычный двойной лист Мебиуса, но в данном случае граница три раза обходит вокруг вертикальной оси. Если трилистник был мал в диаметре, то эта поверхность реализуется как устойчивая мыльная пленка, натянутая на граничный контур A .

Однако X не является ни $\mathbb{Z}$ -поверхностью, ни $\mathbb{Z}_2$ -поверхностью с границей A . Это следует из того, что X гомотопически эквивалентно окружности и гомоморфизм $i_* : H_1(A, \mathbb{Z}) \rightarrow H_1(X, \mathbb{Z})$ переводит образующую 1 группы $\mathbb{Z}$ в элемент $3 \in \mathbb{Z}$. Следовательно, i_* нетривиален над $\mathbb{Z}$ и над $\mathbb{Z}_2$. Однако он тривиален для группы $\mathbb{Z}_3$. Следовательно, тройной лист Мебиуса является $\mathbb{Z}_3$ -поверхностью с границей окружность.

Таким образом, варьируя группы коэффициентов G , мы можем представлять различные пары пространств (X, A) в виде G -поверхностей X с границей A . Интересно, что в примере 3 граничная окружность A , три раза обходящая вокруг оси, является в то же время границей и обычного листа Мебиуса

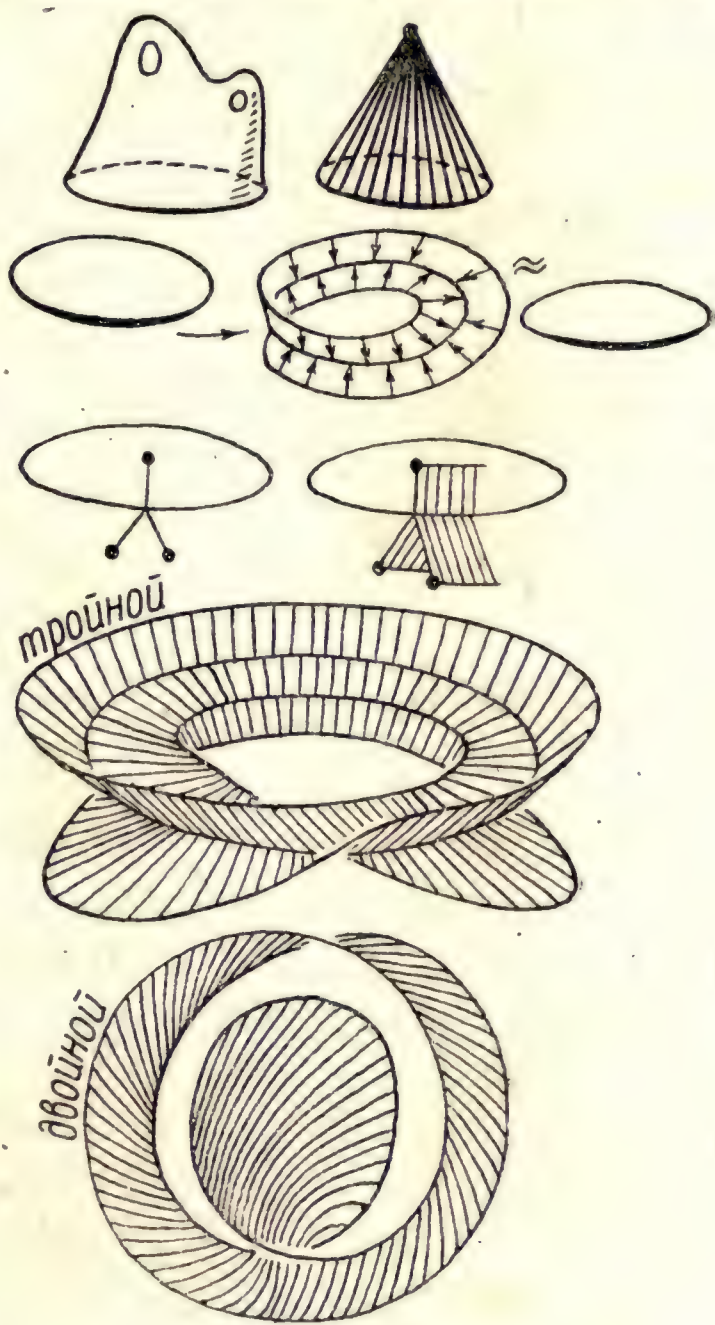


Рис. 47

(рис. 47). Эта поверхность уже не имеет особенностей и также реализуется устойчивой мыльной пленкой.

Если мы не заинтересованы в разделении ориентируемого и неориентируемого случаев, то использование группы $\mathbb{Z}_2$ наиболее оправдано в достаточно широком классе задач. Одним из основных результатов, доказанных в этом направлении, была теорема Райфенберга для случая компактных поверхностей в $\mathbb{R}^n$, расширенная затем Морри на случай произвольных гладких римановых многообразий. Мы сформулируем ее сразу в общем случае. Пусть M — гладкое полное риманово многообразие и $A \subset M$ — компактное измеримое $(k-1)$ -мерное подмногообразие. Пусть G — компактная абелева группа и пусть $H_{k-1}(A, G) \neq 0$. Пусть существует хотя бы одно измеримое k -мерное подмногообразие X , вложенное в M и являющееся G -поверхностью с границей A . Рассмотрим класс $\{X\}$ всех таких поверхностей и положим $d = \inf_{X \in \{X\}} \text{vol}_k(X \setminus A)$, т. е. d яв-

ляется точной нижней гранью объемов всех G -поверхностей с границей A . Тогда утверждается, что $d > 0$ и существует минимальная поверхность X_0 из класса $\{X\}$ такая, что $d = \text{vol}_k(X_0 \setminus A)$. За исключением множества S точек k -мерной меры нуль эта поверхность $X_0 \setminus A$ является открытым гладким минимальным подмногообразием в смысле дифференциальной геометрии, т. е. его средняя кривизна равна нулю. Если объемлющее многообразие M аналитическое, то и подмногообразие $X_0 \setminus (A \cup S)$ также является аналитическим.

Такие поверхности, реализующие абсолютный минимум объема, естественно назвать глобально минимальными.

3. Удивительные примеры физических устойчивых минимальных поверхностей, ретрагирующихся тем не менее на свою границу.

Основным ограничением, сужающим область применимости этой замечательной теоремы, является тот факт, что существуют естественные мыльные пленки, заклеивающие «хорошую границу» A , однако не являющиеся G -поверхностями с границей A ни для какой абелевой группы G . Другими словами, предлагаемое алгебраическое определение границы не охватывает некоторые геометрические пленки, для которых существует естественная физическая реализация. Это обстоятельство мы проиллюстрируем ниже интересным примером Дж. Ф. Адамса.

Предварительно рассмотрим более подробно минимальные пленки, затягивающие окружности, обходящие два и три раза вокруг некоторой оси. Рассмотрим контур, показанный на рис. 48(а). Ясно, что если две петли достаточно близки друг к другу, то на этот контур натягивается мыльная пленка, диффеоморфная обычному листу Мебиуса (рис. 48(б)). Эта пленка устойчива и неориентируема, т. е. является $\mathbb{Z}_2$ -поверхностью с границей окружность. Однако наряду с этой пленкой имеется еще одна устойчивая пленка с той же границей, но сущест-

венно отличающаяся от листа Мебиуса. Если две петли достаточно далеки друг от друга, то существует мыльная пленка, гомеоморфная диску и затягивающая этот контур (рис. 48(в)). Если мы сблизим два витка контура, то диск не может удержаться в прежнем положении, отдельные его части начнут сближаться и, наконец, пленка коллапсирует на новую поверх-

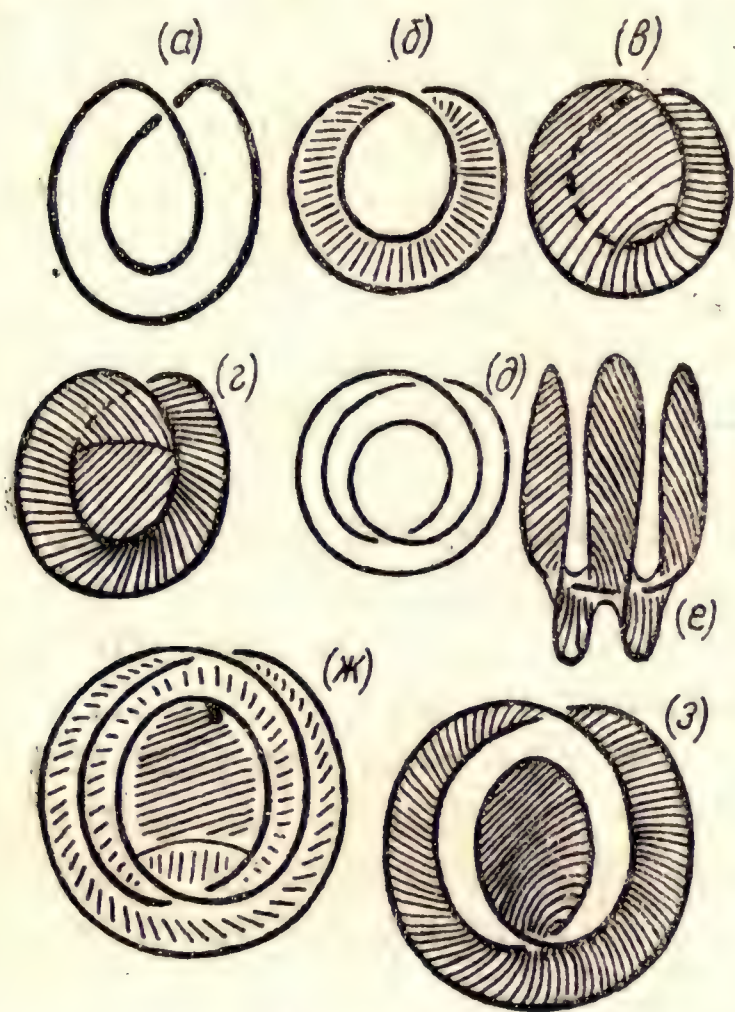


Рис. 48(а) — (з)

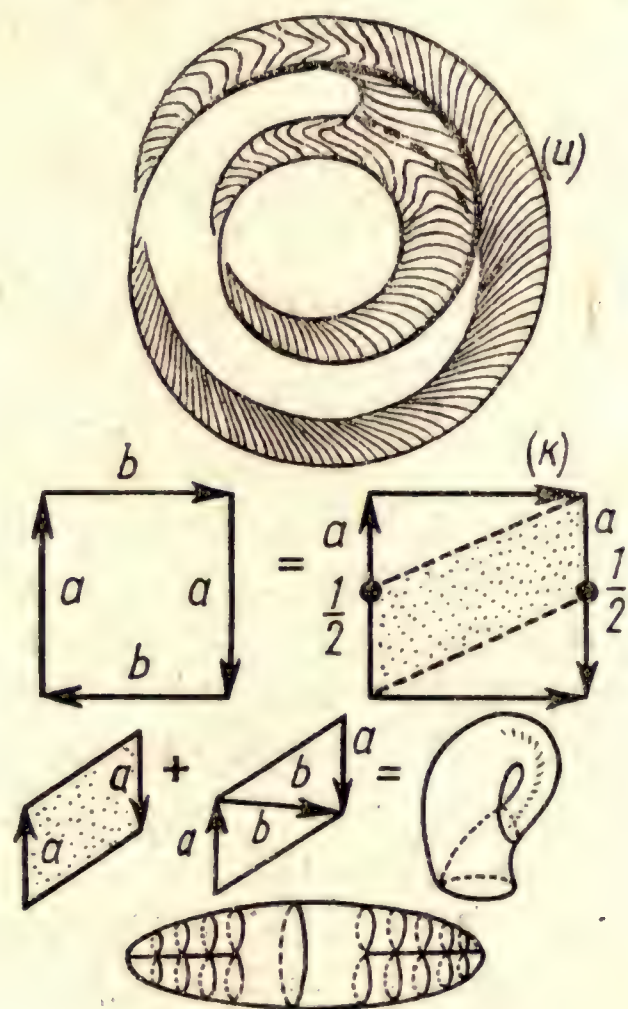


Рис. 48(и), (к)

ность (рис. 48(г)). Эта пленка имеет особое ребро, состоящее из тройных сингулярных точек. Аналогичные события происходят и для контура, делающего три оборота вокруг оси (рис. 48(д)). Если три витка развести достаточно далеко, то на этот контур натягивается устойчивая мыльная пленка, гомеоморфная диску (рис. 48(е)). Ясно, что как только мы начнем сближать три витка окружности, придавая им положение, показанное на рис. 48(д), в конце концов наступит момент, когда диск потеряет устойчивость и коллапсирует на новую минимальную поверхность, уже не гомеоморфную диску (рис. 48(ж)). Отметим, что существуют еще две устойчивые неориентируемые поверхности, имеющие ту же границу, что и тройной лист Мебиуса. Одна из них — это обычный лист Мебиуса (рис. 48(з)), другая получается из бутылки Клейна выбрасыванием диска (рис. 48(и)). Напомним, что бутылка Клейна получается склейкой двух листов Мебиуса по их общей границе (рис. 48(к)). Ясно, что поверхность, изображенная на рис. 48(и), получается склейкой двух листов Мебиуса (наружного и внутреннего — меньшего диаметра) и выбрасыванием диска.

Перейдем к примеру Адамса. Рассмотрим контур, являющийся границей тройного листа Мебиуса (рис. 48(д)). Мы уже знаем, что существует минимальная поверхность с этой грани-

цей для группы Z_2 (рис. 48(ж)). Здесь средняя дырка, образуемая контуром, затянута диском. Другая мыльная пленка — тройной лист Мебиуса (рис. 47) — уже не является Z_2 -поверхностью, но зато является Z_3 -поверхностью. Казалось бы, подбирая каждый раз подходящую группу коэффициентов G , мы можем представить в виде G -поверхности любую «разумную пленку» с данной границей. Однако это не так. Рассмотрим контур A , показанный на рис. 49(а) и являющийся границей устойчивой мыльной пленки X , показанной на рис. 49(б) и получающейся, если мы склеим двойной и тройной листы Мебиуса, соединив их тонкой полоской-мостиком. Очевидно, что эта пленка реализуется физической мыльной пленкой и является вполне «разумной поверхностью» с данной границей. Границей пленки является окружность, вложенная без самопересечений в R^3 .

Однако ни для какой абелевой группы G эта пленка не является G -поверхностью с границей окружность. Дело в том, что пленка ретрагируется на свою границу! Напомним, что подпространство A пространства X называется ретрактом X , если существует непрерывное отображение $f: X \rightarrow A$, определенное на всем X и оставляющее на месте все точки из A . Другими словами, пленка «мгновенно» коллапсирует на свою границу, остающуюся при этом неподвижной. Доказательство ретрагируемости мы дадим ниже, а здесь извлечем отсюда необходимое следствие. Пусть $i: A \rightarrow X$ — вложение, а $f: X \rightarrow A$ — ретракция. Сквозное отображение $A \xrightarrow{i} X \xrightarrow{f} A$ совпадает с тождественным. Сквозной гомоморфизм

$$H_1(A, G) \xrightarrow{i_*} H_1(X, G) \xrightarrow{f_*} H_1(A, G)$$

является тождественным, и, в частности, i_* — мономорфизм. Следовательно, для любой группы G пленка X не заклеивает ни одного цикла в $H_1(A, G)$, т. е. X не является G -поверхностью с границей A . Более того, гомоморфизм $i_*: F(A) \rightarrow F(X)$ является мономорфизмом для любого естественного «функтора» F , например для гомотопических групп. Следовательно, A не является границей X ни в каком разумном алгебраическом смысле, хотя X является устойчивой пленкой и реализует абсолютный минимум площади в классе всех «разумных пленок» с границей A .

Докажем, что A — ретракт X , воспользовавшись теоремой Хопфа (гл. 1, § 2). В качестве A возьмем S^1 , пусть X — сумма тройного и двойного листов Мебиуса. Вычислим гомоморфизм $i_*: H_1(A, U) \rightarrow H_1(X, U)$. Так как пространство X гомотопически эквивалентно букету двух окружностей, то $H_1(X, U) = U \oplus U$ и $i_*(\lambda) = (2\lambda, 3\lambda)$, где $0 \leq \lambda \leq 1$. Отсюда видно, что i_* — мономорфизм, что и требовалось доказать. Отметим, что хотя A — ретракт X , но A не является деформационным ретрактом X . Напомним, что подпространство $A \subset X$ является деформационным ретрактом, если существует непрерывная гомотопия

$\varphi_t: X \rightarrow X$ такая, что φ_0 — тождественное отображение X на себя, а φ_1 отображает X в A , причем для всех t отображение φ_t оставляет A неподвижным. Так как композиция $A \xrightarrow{i} X \xrightarrow{\varphi_1} A$ является тождественным отображением, а композиция $X \xrightarrow{\varphi_1} A \xrightarrow{i} X$ гомотопна тождественному отображению, то X и A гомотопически эквивалентны. В нашем случае A — окруж-

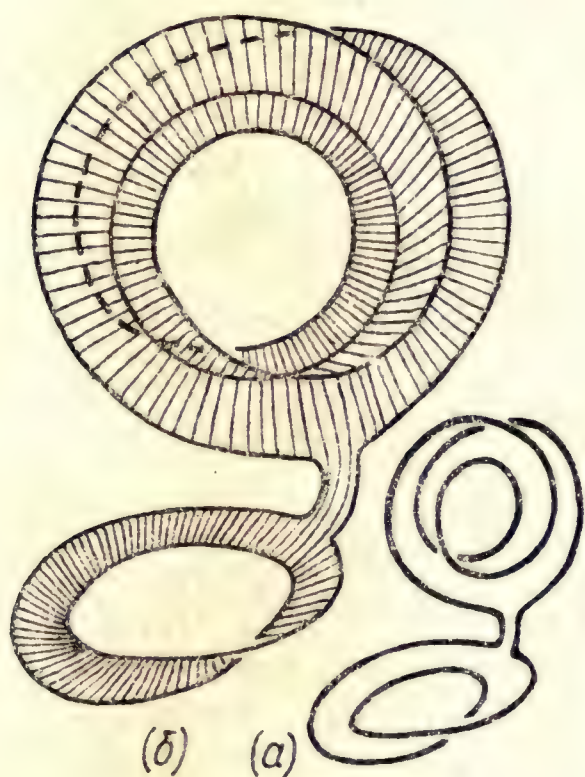


Рис. 49

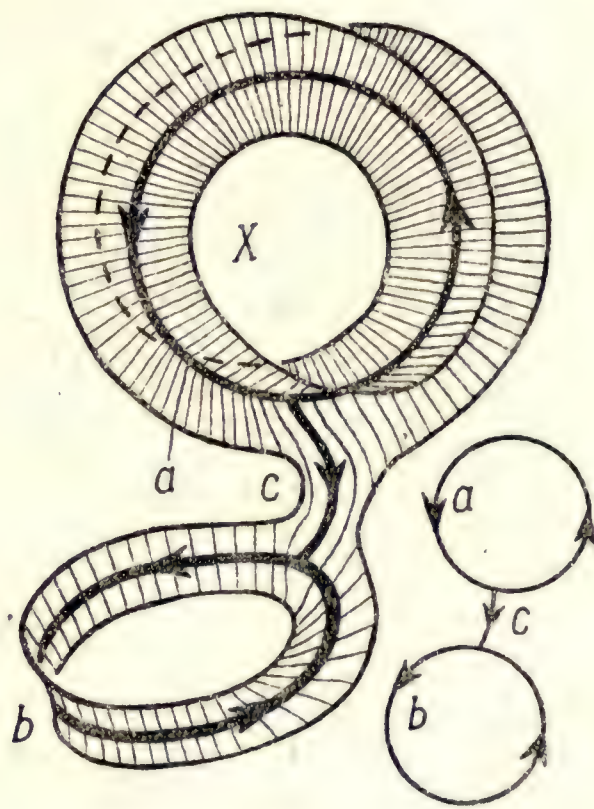


Рис. 50(a)

ность, X эквивалентно букету двух окружностей, т. е. A и X гомотопически не эквивалентны, т. е. A не является деформационным ретрактом.

Существование ретракции было получено из алгебраических свойств этих комплексов. Полезно наглядно представить, как происходит эта ретракция геометрически. Это построение выполнено Т. Н. Фоменко. Воспользуемся приведенным в гл. 1 гео-

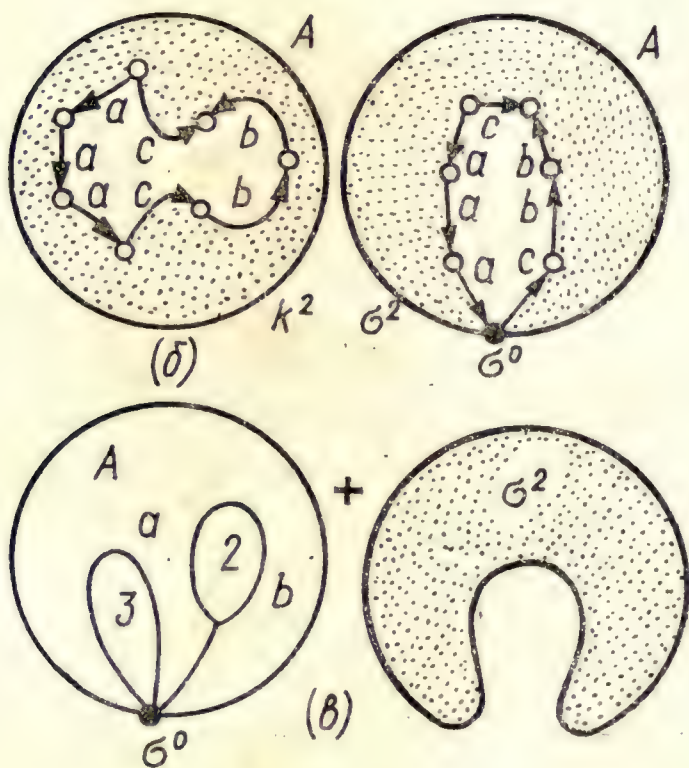


Рис. 50(б), (в)

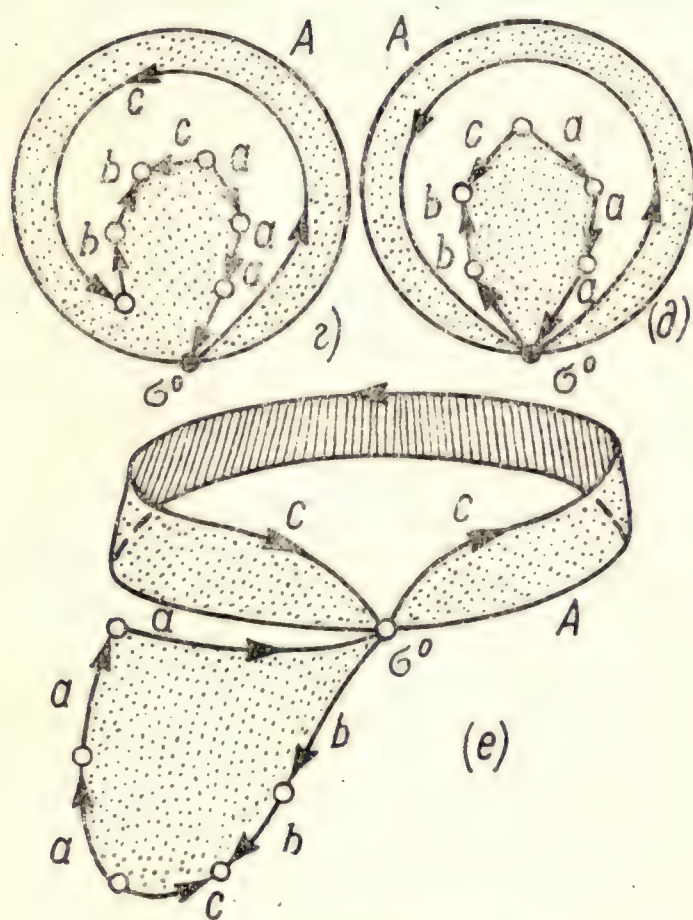


Рис. 50(г) — (е)

метрическим доказательством теоремы Хопфа. В нашем случае комплекс X двумерен, и его одномерный остов совпадает с букетом двух окружностей, одна из которых является осью обычного листа Мебиуса, а другая — осью тройного листа Мебиуса. В последнем случае окружность состоит из сингулярных точек мыльной пленки. Поступая по схеме § 2, гл. 1, рассмотрим клеточное разбиение X . Ясно, что $X = \sigma^0 \cup \sigma_1^1 \cup \sigma_2^1 \cup \sigma^2$, где σ^2 — единственная двумерная клетка. Как приклеивается она к одномерному остову? Разрежем мыльную пленку вдоль осей b и a

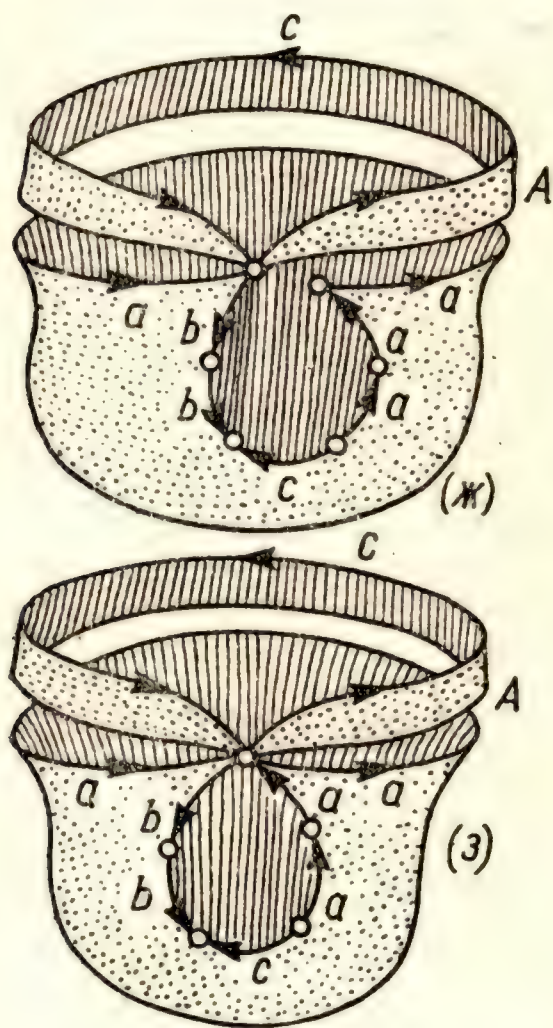


Рис. 50(ж), (з)

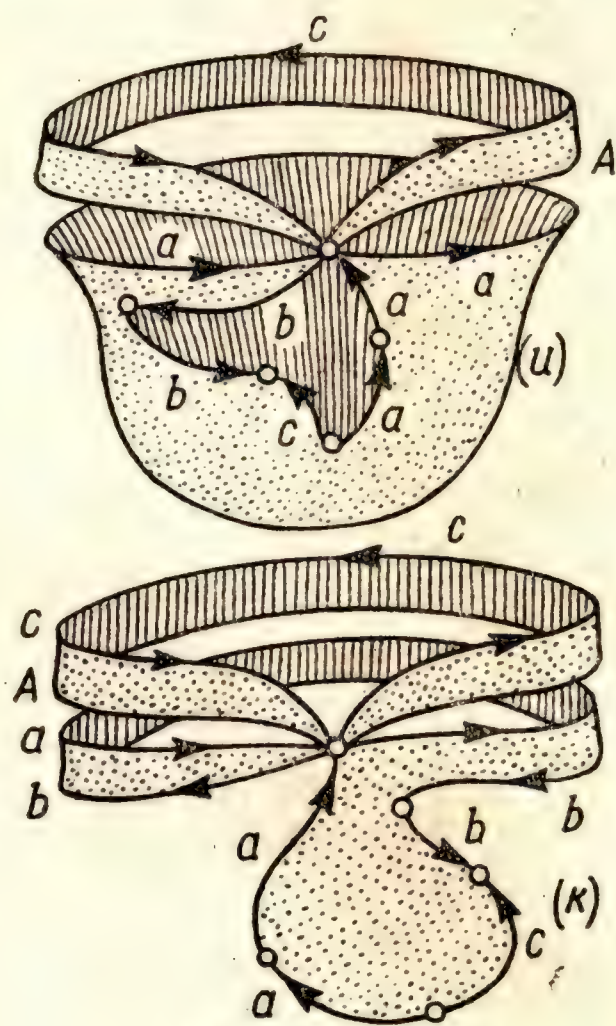


Рис. 50(и), (к)

соответственно двойного и тройного листов Мебиуса и вдоль короткого отрезка c , являющегося осью перепонки-мостика, соединяющего эти два листа Мебиуса (рис. 50(а)). Пленка X представляется в виде: $X = \sigma^0 \cup a \cup b \cup c \cup k^2$, где k^2 — кольцо. После разреза по одномерному остову пленка превращается в плоскую область, гомеоморфную обычному кольцу k^2 , граница которого состоит из двух окружностей: внешней, являющейся исходной границей A пленки X , и внутренней, представленной в виде объединения семи дуг: трех экземпляров отрезка a , двух экземпляров отрезка b и двух экземпляров отрезка c (рис. 50(б)). Для удобства осуществим деформационную ретракцию, в результате которой внутренняя граница соприкоснется в одной своей точке с внешней границей (рис. 50(в)). Получившаяся полная граница двумерной области ограничивает двумерную клетку σ^2 . Для установления соответствия с обозначениями теоремы 3 из § 2 гл. 1 обозначим A через S_0^1 (это выделенная окружность в одномерном остове), окружность a — через S_1^1 и окружность b — через S_2^1 . Тогда коэффициенты

инцидентности клетки σ^2 с этими окружностями таковы: $[o_2 : A] = 1$, $[\sigma^2 : a] = 3$, $[\sigma^2 : b] = 2$. Следуя конструкции § 2 гл. 1, нужно построить отображение g одномерного остова пленки X на окружность A . Отообразим A на себя тождественно. Самое главное — определить g на окружностях a и b . Нужно подобрать такие целые числа m и n , чтобы $m \cdot [\sigma^2 : a] - n \cdot [\sigma^2 : b] = 1$, т. е. $3m - 2n = 1$. Затем следует построить отображение g , которое отобразит a со степенью m на A , а b со степенью n на A .

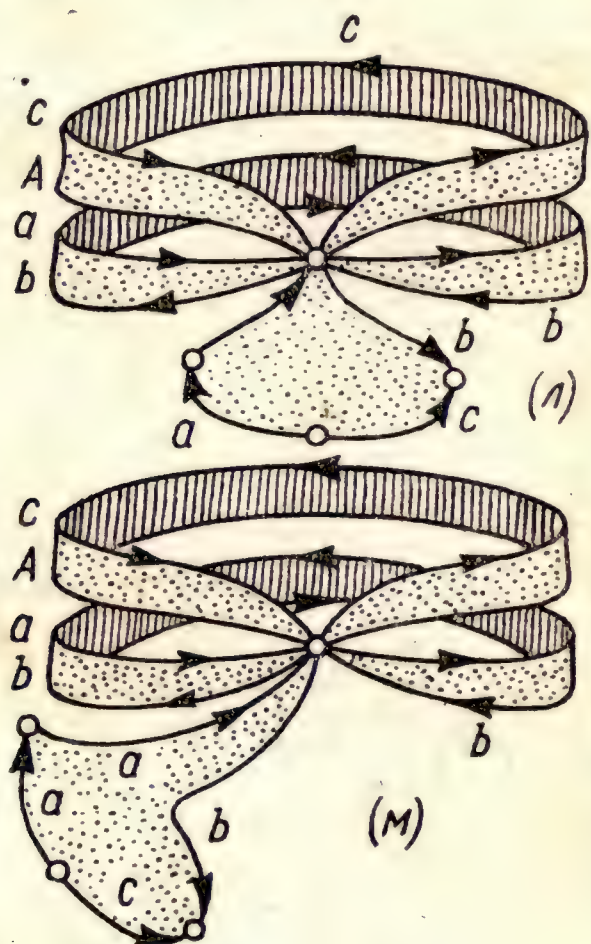


Рис. 50(л), (м)

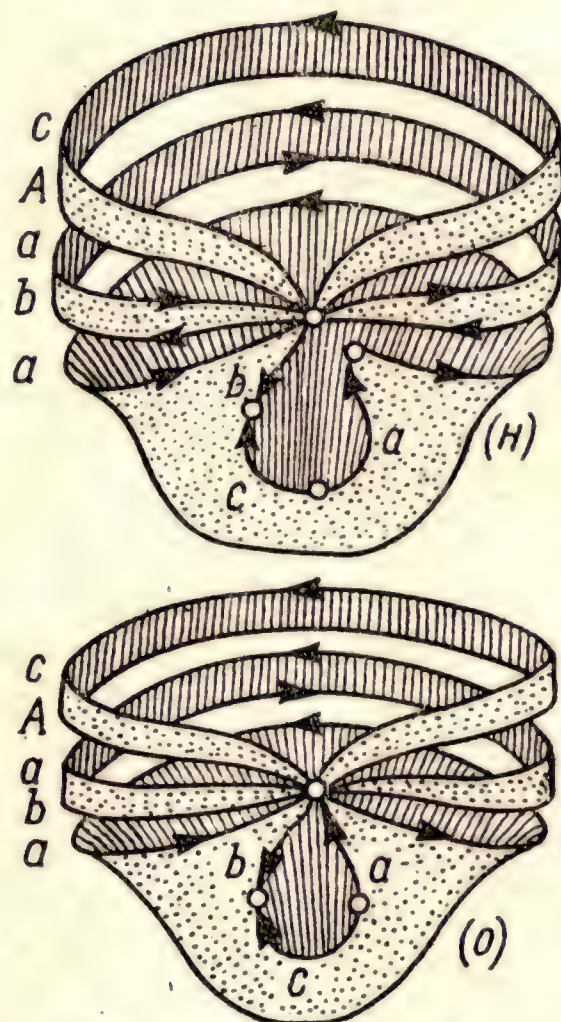


Рис. 50(н), (о)

В нашем случае достаточно положить $m = n = 1$, т. е. можно считать, что a и b отображаются на A гомеоморфно. Согласно теореме 3 (§ 2 гл. 1) это отображение g продолжается до отображения f всей пленки на A , что и требуется. Осталось изобразить это отображение наглядно. Оказывается, после того как мы разрезали пленку по одномерному остову, искомое отображение f реализуется в $\mathbb{R}^3$ как последовательность непрерывных деформаций клетки σ^2 . Начнем с того, что фиксируем A и растянем сторону c так, чтобы она намоталась на A со степенью единица (рис. 50(г)). Но при этом мы не будем отождествлять A и c и прервем операцию, сохранив узкую полоску, ограниченную на рис. 50(д) окружностями A и c . Оставшаяся часть двумерной клетки образует диск, приклеенный к точке σ^0 в получившейся полоске между A и c . Берем ребро a и наматываем его вдоль A так, как показано на рис. 50(е) и рис. 50(ж). Свисающий с точки σ^0 диск обворачивается вокруг A , и в результате получается мешок с границей a , причем в мешке имеется дырка с границей $bbs^{-1}aa$. Это слово выписано при движении против часовой стрелки на гра-

нице дырки, степень $+1$ или -1 указывает на ориентацию соответствующего ребра (рис. 50(з)). Продолжаем намотку двумерной клетки. Берем b и также наматываем его вдоль A , действуя так, как показано на рис. 50(и). В результате получается картина, показанная на рис. 50(к), и итогом этой операции является поверхность, изображенная на рис. 50(л). Граница свисающего диска упростилась по сравнению с предыду-

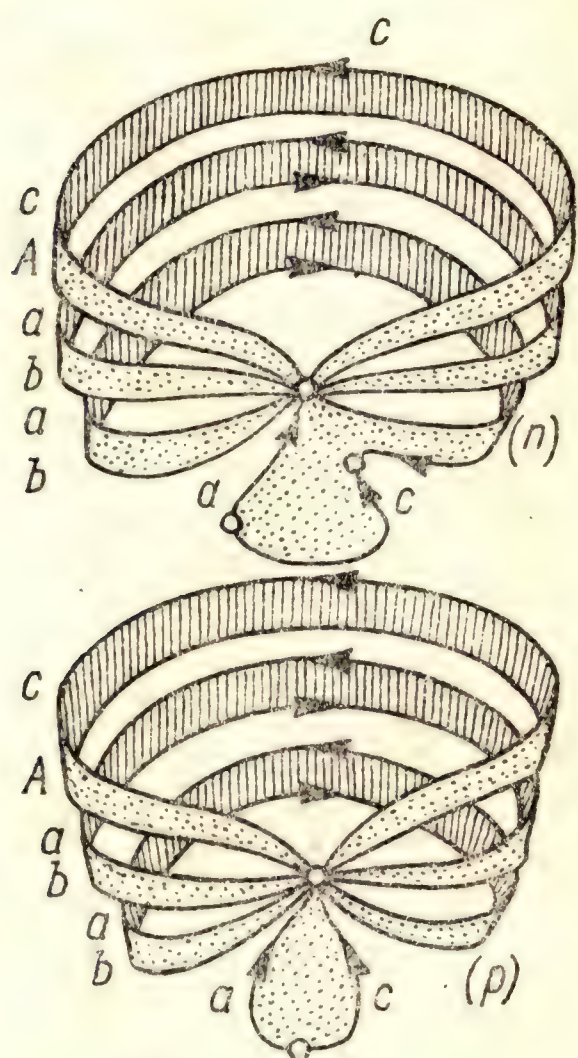


Рис. 50(п), (р)

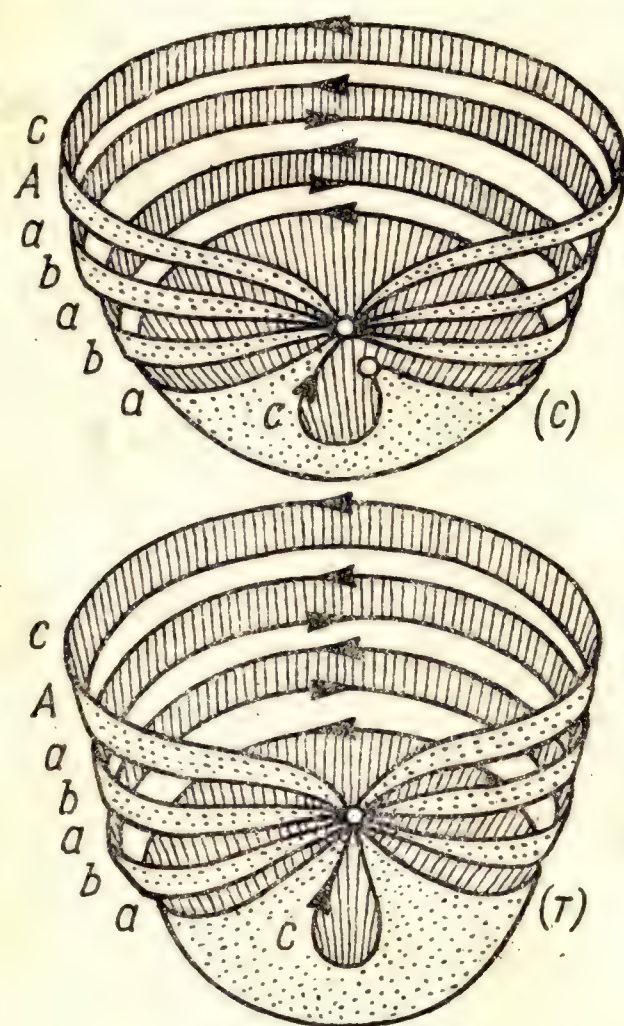


Рис. 50(с), (т)

щим шагом и имеет вид $a^{-1}a^{-1}cb^{-1}$. Берем следующий экземпляр ребра a и наматываем его вдоль A , как показано на рис. 50(м) и 50(н). Итак, мы последовательно чередуем намотку ребер типа a и типа b . Результат см. на рис. 50(о). Мы снова получили мешок с границей a и с дыркой bca . Берем оставшийся экземпляр ребра b (на этот раз он последний) и наматываем его вдоль A как показано на рис. 50(п). В итоге получаем поверхность на рис. 50(р), содержащую свисающий диск с границей $a^{-1}c$. Берем последний экземпляр ребра a и снова повторяем намотку, показанную на рис. 50(с). В результате получаем поверхность на рис. 50(т) с мешком, границей которого является a и который содержит дырку с границей c . На последнем шаге растягиваем сторону c , раздуваем дырку и добиваемся того, что мешок превращается в узкую полосу, сдвигающую A (рис. 50(у)). Построенная нами непрерывная деформация разрезанной пленки X замечательна тем, что теперь пленка состоит из четырех узких полосок, соединенных в одной точке σ^0 и расположенных в пространстве вдоль окружности A , в непосредственной близости от нее. Теперь все ребра

a ориентированы в одном направлении относительно A и все ребра b также ориентированы одинаково, в направлении, противоположном направлению ребер a . Два ребра c также ориентированы одинаково (в направлении a). Так как все ребра одного типа ориентированы одинаково, то переходим к заключительному этапу. Стягиваем каждую из четырех полосок на ее ось, среднюю линию, и получаем четыре окружности на

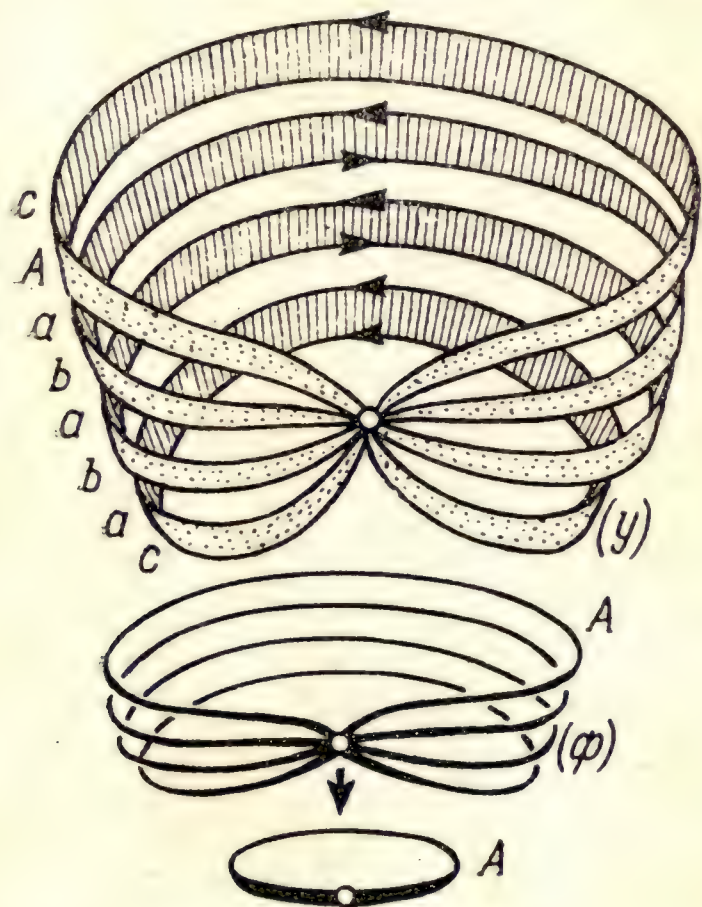


Рис. 50(y), (ф)

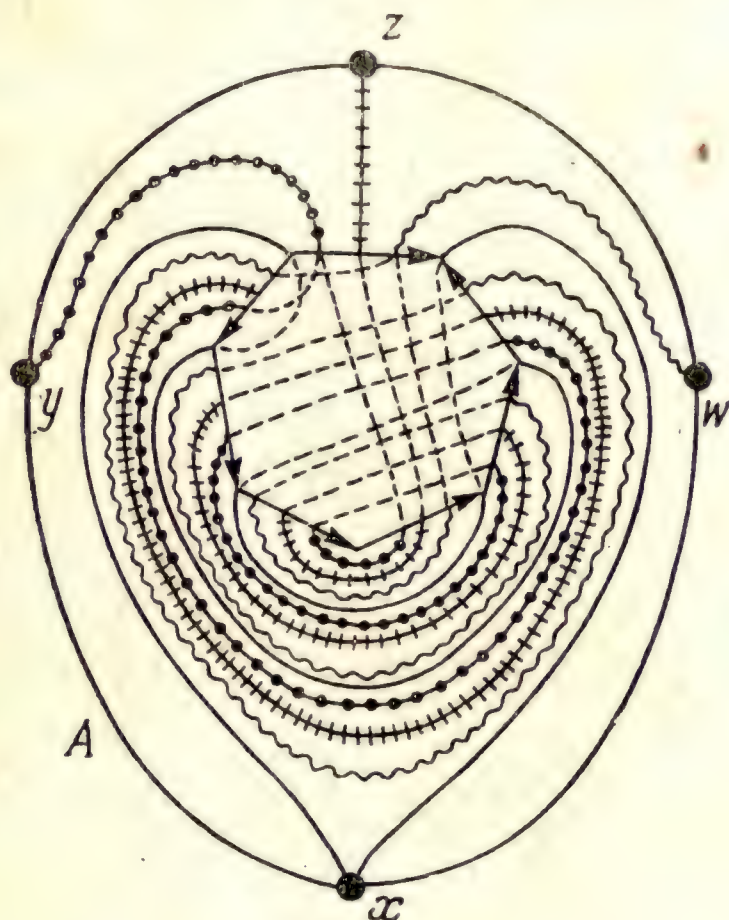


Рис. 50(x)

рис. 50(ф), которые одним и тем же способом отображаем на A . Построенное отображение непрерывно вследствие согласованности всех ребер одного типа на поверхности рис. 50(y). Этим и завершается построение искомой ретракции. Тот же механизм действует и в случае, когда мы рассматриваем сумму не двойного и тройного листов Мебиуса, а сумму p -кратного и q -кратного листов Мебиуса, где числа p и q взаимно-просты. Наглядное изображение ретракции в принципе сохраняется, однако добавляется намотка окружностей a и b со степенями m и n на окружность A , где числа m и n таковы, что $mp - nq = 1$. Эта намотка усложняет геометрическую картину, но не меняет принципа построения ретракции.

В заключение приведем чертеж, показывающий, как распределены на плоской области k^2 (кольцо с границей $A + 3a + 2b$) прообразы точек A . Рассмотрим точку x из A , и пусть f — ретракция X на A . Рассмотрим полный прообраз $f^{-1}(x)$. Он является системой непрерывных дуг в кольце k^2 . На рис. 50(x) выбраны 4 точки x, y, z, w на A , и для каждой из них нарисован полный прообраз. Он состоит, вообще говоря, из нескольких компонент связности. Пунктир показывает склейку траекторий, принадлежащих одному прообразу.

Приведем также второй пример Адамса — минимальную пленку X , затягивающую контур A , показанный на рис. 51(a).

Этот контур гомеоморфен окружности, вложенной без самопересечения в $\mathbb{R}^3$. Пленка, показанная на рис. 51(б), очевидно, является «хорошей минимальной пленкой», правда, неустойчивой и имеющей особенности. Однако, оказывается, что пленка не является G -поверхностью с границей A ни для какой группы

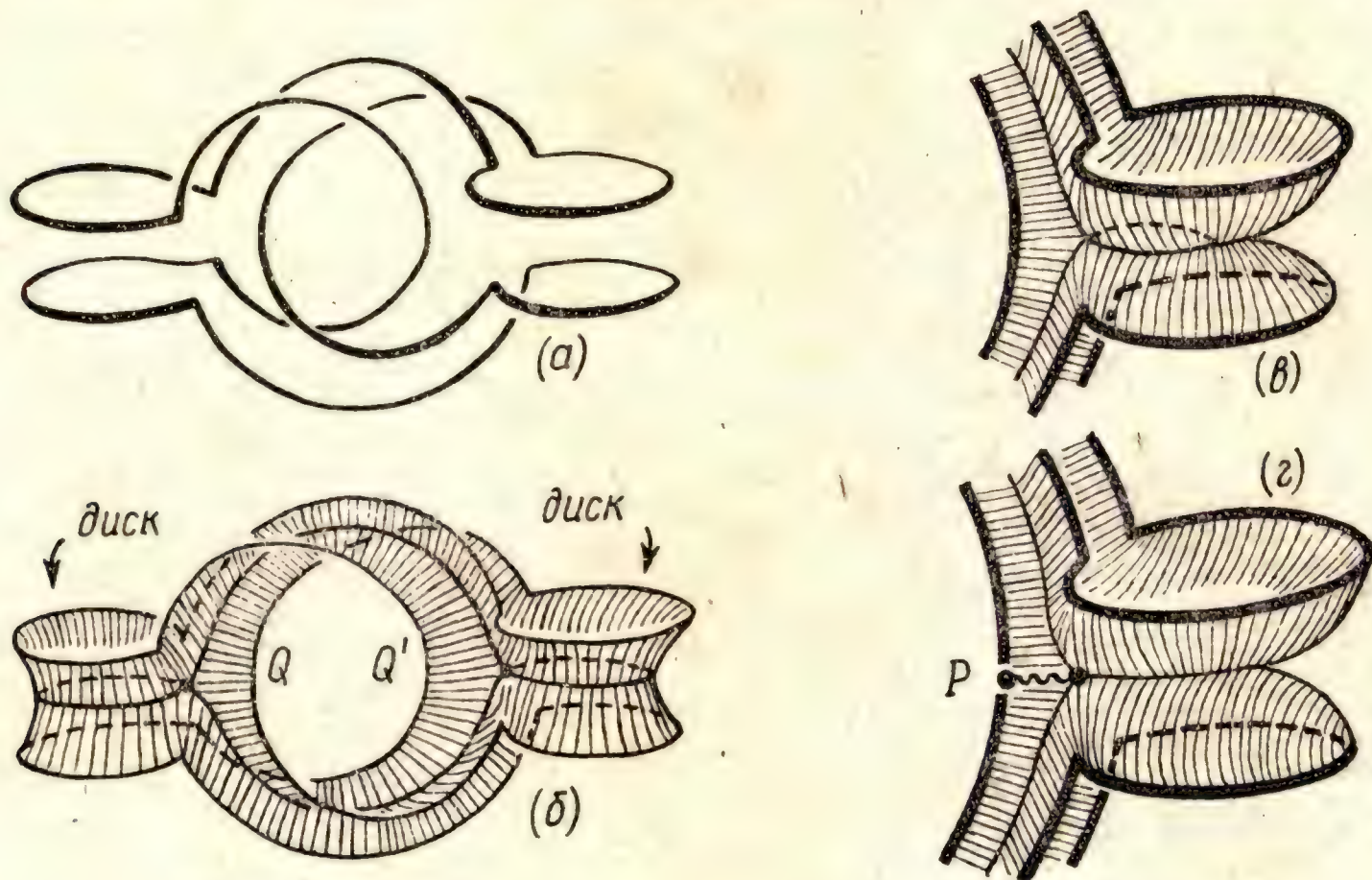


Рис. 51(а), (б)

Рис. 51(в), (г)

G , поскольку она ретрагируется на свою границу. Более того, существует не только обычная ретракция, но и деформационная ретракция X на A . Непрерывное стягивание пленки на ее границу, оставляющее границу все время неподвижной, см. на рис. 51 [19]. Это означает, что разрушение указанной неустойчивой пленки происходит не скачком, как в первом примере, а постепенно: пленка непрерывно деформируется по себе, сокращается и, наконец, стягивается на свою границу. Две ручки справа и слева (рис. 51(б)) образованы парами соосных окружностей, прикрепленных к основному контуру. Поскольку окружности близки друг к другу, на них натягивается мыльная пленка, напоминающая катеноид. Горловина каждого из них затянута плоским диском, подклеенным изнутри к окружности наименьшего радиуса на катеноиде. Особое внимание обратим на две сингулярные точки Q и Q' , в каждой из которых сходится по 4 сингулярных ребра пленки под углами примерно в 109° . Кроме того, в каждую из точек входит по 6 листов пленки. Около точек Q, Q' пленка устроена как в окрестности центра тетраэдральной пленки (рис. 39). Чтобы изобразить деформационную ретракцию, полезно слегка утолщить пленку, что позволяет более свободно обращаться со скольжением пленки по себе. Когда какая-то часть пленки передвигается по себе, удобно, используя «толщину пленки», изображать эти части раздельно. Напомним, что деформация будет протекать внутри утолщенной пленки. На первом шаге стянем по

себе в точку каждый из плоских дисков, заклеивающих горловины боковых катеноидов (рис. 51(в), (г)). Можно считать, что диски втягиваются в точки Q и Q' . Следовательно, два выпуклых диска, заклеивающих верхнюю и нижнюю окружности на отдельной ручке, например, справа, можно считать не со-

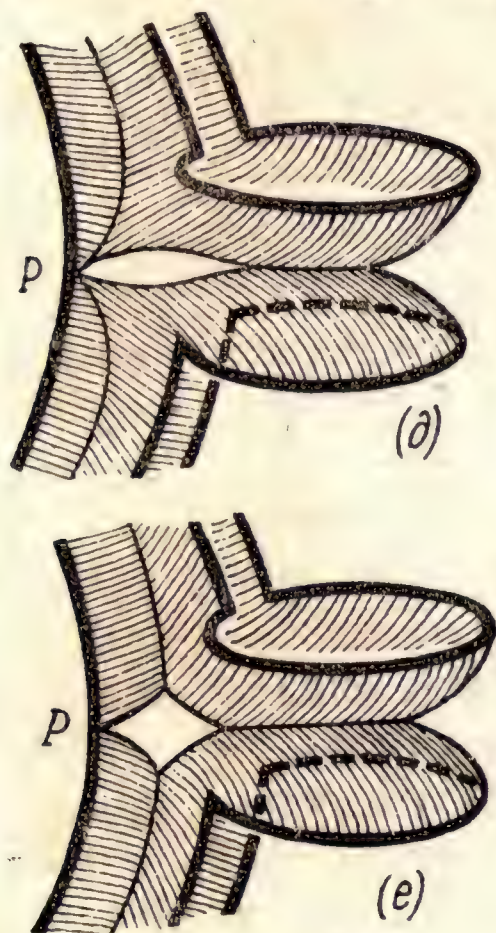


Рис. 51(д), (е)

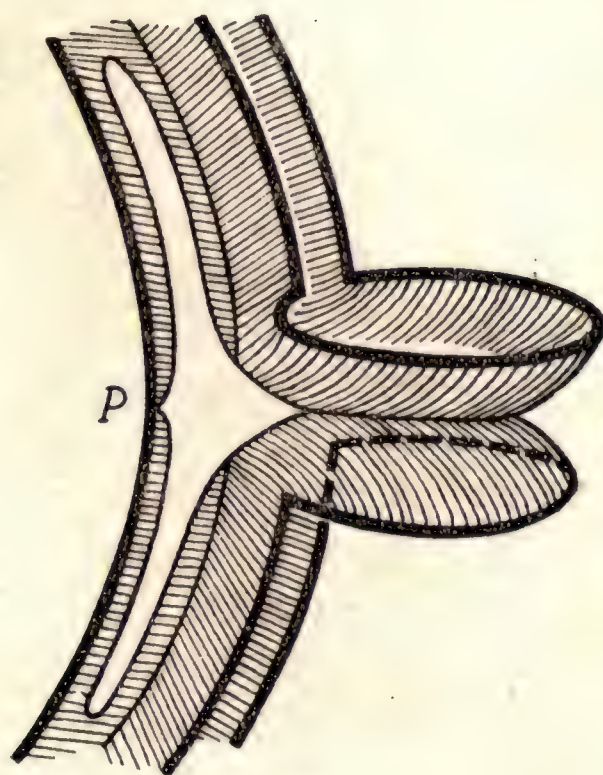


Рис. 51(ж)

прикасающимися внутри утолщенной пленки. Поэтому отрезок c (рис. 51(г)) можно стянуть в точку P , и возникает ситуация, показанная на рис. 51(д). Рисунки 51(д) и 51(е) отличаются лишь тем, что сингулярные тройные ребра заняли положения, соответствующие стягиванию реальной мыльной пленки. Получившаяся свободная зона, дырка в пленке, начинает неудержимо распространяться вдоль пленки, монотонно уменьшая ее площадь (рис. 51(ж), (з), (и)). Из каждой ручки вытягивается язык, который съедает свою часть поверхности, двигаясь вдоль бывшей окружности тройных особых точек. Наконец эти языки возвращаются к двум парам боковых дисков, съедают и их, что завершает процесс стягивания. При всей удивительности свойств этих пленок уместно напомнить, что мы уже сталкивались с не менее любопытными (хотя и более простыми) примерами реальных пленок, затягивающих незамкнутые контуры (рис. 46).

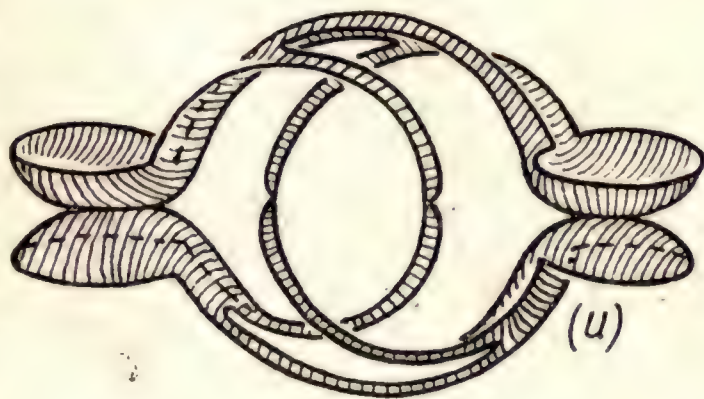
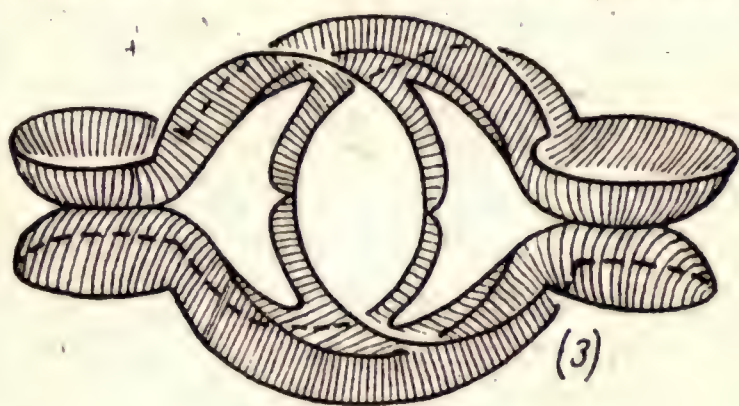


Рис. 51(з), (и)

4. Когда мыльная пленка, затягивающая контур, не содержит в себе замкнутых мыльных пузырей?

Перейдем к изучению глобальных топологических свойств минимальных поверхностей X размерности k , заклеивающих в алгебраическом, гомологическом смысле $(k-1)$ -мерные «контур», вложенные в некоторое риманово многообразие M . Как мы уже знаем, если $H_{k-1}(A, G) \neq 0$ и если существует хотя бы одна пленка, затягивающая A , то всегда существует минимальная поверхность X_0 с границей A , реализующая абсолютный минимум функционала объема (площади) в классе всех G -поверхностей с границей A . В реальных опытах мы часто получаем пленки, содержащие в себе как куски поверхностей нулевой средней кривизны, так и мыльные пузыри, ограничивающие замкнутые объемы. Возникает вопрос: (а) в каких случаях мыльная пленка состоит только из поверхностей нулевой средней кривизны; (б) в каких случаях она включает поверхности постоянной положительной средней кривизны — пузыри. Особый интерес представляют «настоящие минимальные пленки», т. е. пленки, целиком состоящие из поверхностей нулевой кривизны, без пузырей. Существуют простые критерии, которым должна удовлетворять граница A для того, чтобы пленка, реализующая абсолютный минимум объема (площади), не содержала пузырей. Эти критерии обнаружены автором книги.

Математическое понятие гомологического цикла является прекрасной моделью для реальных мыльных пузырей. В некоторых случаях можно представлять себе, что цикл является замкнутой, т. е. не имеющей края, гладкой поверхностью. Гомологические циклы размерности k естественно образуют k -мерную группу гомологий пространства X_0 . Поэтому с топологической точки зрения вопрос о том, содержит ли k -мерная поверхность X_0 «пузыри» или нет, допускает следующую переформулировку: содержит ли поверхность X_0 k -мерные циклы (для данной группы коэффициентов)? Другими словами, k -мерная поверхность X_0 не содержит «пузырей» в том и только в том случае, когда $H_k(X_0, G) = 0$. Например, все три мыльные пленки, показанные на рис. 44, содержат в себе замкнутые пузыри и имеют, следовательно, нетривиальную группу двумерных гомологий, изоморфную группе коэффициентов G , т. е. группа $H_2(X_0, G)$ имеет одну образующую. Таким же свойством обладают и соответствующие этим пленкам живые радиолярии (рис. 44). Напротив, пленки, показанные, например, на рис. 39, 49, 51, 48, 38, имеют нулевую группу двумерных гомологий. При этом некоторые из этих пленок весьма нетривиальны и, в частности, не стягиваются по себе в точку. Итак, для выяснения того, содержит ли минимальная поверхность X_0 «пузыри» — циклы, нужно вычислить $H_k(X_0, G)$.

Предложение 1 (см. [13, 15]). Пусть M — полное гладкое риманово многообразие и $A \subset M$ — компактное измеримое $(k-1)$ -мерное подмножество, например, $(k-1)$ -мерное подмно-

многообразии в M . Пусть G — абелева группа, являющаяся одномерным векторным пространством над некоторым полем F . Например, G — группа $\mathbb{R}^1$ вещественных чисел, или $G = \mathbb{Z}_p$, где p — простое число. Пусть существует хотя бы одна G -поверхность в M с границей A . Тогда глобально минимальная G -поверхность X_0 с границей A обладает тем свойством, что $H_k(X_0, G) = 0$, т. е. не содержит k -мерных циклов — «пузырей».

Доказательство см. ниже. Если объемлющее многообразие M устроено более сложно, чем евклидово пространство, то оно само может содержать нетривиальные k -мерные циклы — «дырки», т. е. группа $H_k(M, G)$ может быть отлична от нуля. В приложениях часто встречаются случаи, когда минимальная пленка, заклеивающая некоторую границу $A \subset M$, реализует в то же время и некоторые k -мерные циклы в M , т. е. обволакивает соответствующие дырки. Поэтому наряду с задачей заклейки границы A возникает не менее естественная геометрическая задача реализации с помощью минимальных поверхностей тех циклов, которые имеются в объемлющем пространстве M . Дадим формальное определение реализации.

Фиксируем в $H_k(M, G)$ какую-либо нетривиальную подгруппу P . Будем говорить, что подмножество $X \subset M$ реализует подгруппу P , т. е. реализует циклы этой подгруппы, если при гомоморфизме $j_* : H_k(X, G) \rightarrow H_k(M, G)$, индуцированном вложением $j : X \rightarrow M$, подгруппа P содержится в образе группы $H_k(X, G)$.

Рассмотрим класс $[P]$ всех компактных подмножеств X , где $A \subset X \subset M$, реализующих подгруппу P . Оказывается, что в этом классе всегда существует глобально минимальная поверхность X_0 , т. е. такая, что ее объем (площадь) $\text{vol}_k(X_0 \setminus A)$ равен $\inf_{Y \in [P]} \text{vol}_k(Y \setminus A)$, (см. [13, 15]).

Из наглядных соображений естественно ожидать, что минимальная поверхность X_0 , реализующая P , не содержит в себе никаких «лишних» k -мерных циклов, а только те, которые обволакивают ненулевые циклы из группы $H_k(M, G)$. При этом не исключено, что поверхность X_0 реализует не только циклы из подгруппы P , но и еще какие-то дополнительные циклы, не принадлежащие P . Другими словами, гомоморфизм $j_* : H_k(X_0, G) \rightarrow H_k(M, G)$ должен быть мономорфизмом. Никакой цикл на пленке X_0 не должен бесследно раствориться в многообразии M , если пленка минимальна. Интуитивно ясно, что если какой-то цикл из пленки X_0 оказался лишним, т. е. стал гомологичным нулю после вложения пленки в многообразие, то часть пленки, соответствующую этому циклу (его носитель), можно выбросить, не влияя на реализуемость циклов из подгруппы P и уменьшая в то же время объем (площадь) пленки. Для минимальной пленки такая процедура невозможна, поэтому пленка не может содержать лишние растворяющиеся циклы. Это предположение оправдывается (А. Т. Фоменко).

Предложение 2 (см. [13, 15]). Пусть выполнены предположения предложения 1. Тогда если минимальная поверхность X_0 реализует нетривиальную подгруппу $P \subset H_k(M, G)$, то гомоморфизм $j_* : H_k(X_0, G) \rightarrow H_k(M, G)$, индуцированный вложением $j : X_0 \rightarrow M$, является мономорфизмом, т. е. на пленке нет «лишних циклов».

Существование глобально минимальных поверхностей X_0 следует из работ [7—8], а в более общей ситуации — из работ автора [13—15]. Перейдем к доказательству предложений 1, 2.

Лемма 1. Пусть X — k -мерное компактное множество в M , Z — компактное подмножество в X , G — абелева группа, являющаяся одномерным векторным пространством над некоторым полем F , и пусть $H_k(Z, G) = 0$. Предположим, что $X \setminus Z$ является открытым топологическим k -мерным подмногообразием, пусть $l \in H_k(X, G)$, $l \neq 0$. Тогда существует открытый k -мерный диск $D \subset X \setminus Z$ такой, что если $i_* : H_k(X \setminus D) \rightarrow H_k(X)$ — гомоморфизм, индуцированный вложением, то: (а) i_* — мономорфизм, (б) $\text{Codim Im } i_* = 1$ в группе $H_k(X)$, (в) $l \notin \text{Im } i_*$.

Доказательство. Рассмотрим коммутативную диаграмму:

$$\begin{array}{ccccccc}
 & & H_{k-1}(X \setminus D) & & & & \\
 & & \uparrow & & & & \\
 & & H_k(S^k) & & & & \\
 & & \uparrow i_* & & & & \\
 0 = H_k(Z) & \rightarrow & H_k(X) & \xrightarrow{\varphi_*} & H_k(X, Z) & \xrightarrow{f_*} & H_k(X/Z) \\
 \uparrow & & \uparrow i_* & & \uparrow \alpha_* & & \uparrow \omega_* \\
 0 = H_k(Z) & \rightarrow & H_k(X \setminus D) & \rightarrow & H_k(X \setminus D, Z) & \rightarrow & H_k((X/Z) \setminus f(D)) \\
 & & \uparrow & & \uparrow \psi_* & & \uparrow f'_* \\
 & & 0 = H_{k+1}(S^k) & & & &
 \end{array}$$

Здесь горизонтальные строки являются фрагментами точных последовательностей пары, D — некоторый открытый диск в $X \setminus Z$, $f : X \rightarrow X/Z$ — естественная проекция, пространство $S^k = X/(X \setminus D)$ гомеоморфно сфере, отображения f_* и f'_* — изоморфизмы. Положим $q = f_* \varphi_* l$, тогда $q \neq 0$, поскольку φ_* — мономорфизм. Таким образом, мы указали ненулевой элемент q в группе $H_k(X/Z)$, порожденной элементом l , а потому если мы разрушим q выбрасыванием диска $f(D)$ из X/Z , то мы автоматически разрушим и элемент l выбрасыванием диска D из X . Пространство X/Z является k -мерным многообразием всюду, за исключением одной точки x — образа множества Z при проекции f . Если предположить, для простоты, что $(X/Z) \setminus x$ — гладкое многообразие, состоящее из конечного числа компонент линейной связности, то, очевидно, существует такой диск $f(D) \subset (X/Z) \setminus x$, выбрасывание которого из X/Z разрушает элемент q , т. е. q не принадлежит образу гомоморфизма $H_k((X/Z) \setminus f(D)) \rightarrow H_k(X/Z)$. Общий случай нуждается в технически бо-

лее сложных рассуждениях с нервами покрытий, используемых при определении групп спектральных гомологий. Мы не будем здесь на этом останавливаться, отсылая заинтересованного читателя к [13—15]. Итак, мы предъявили диск $f(D) \subset X/Z$, удаление которого разрушает q . Докажем, что l не принадлежит образу гомоморфизма i_* , т. е. что удаление диска D из X разрушает l . Ясно, что $D \subset X \setminus Z$ (рис. 52). Допустим противное, пусть $l = i_*(\alpha)$, где $\alpha \in H_k(X \setminus D)$. Но в силу коммутативности диаграммы имеем: $f_* \varphi_* l = f_* \varphi_* i_* \alpha = \omega_* f'_* \psi_* \alpha$, где $\rho = f'_* \psi_* \alpha \in H_k((X/Z) \setminus f(D))$. Отсюда $q = \omega_* \rho$, что противоречит выбору диска $f(D)$, согласно которому q не принадлежит образу ω_* . Полученное противоречие доказывает пункт (в) леммы. Мономорфность i_* следует из рассмотрения точного вертикального столбца диаграммы, так как $H_{k+1}(S^k) = 0$. Так как $l \neq 0$ из $H_k(X)$ не принадлежит образу i_* , то гомоморфизм j_* не является тождественно-тривиальным, так как в силу точности $j_* l \neq 0$. Но в таком случае образ j_* покрывает всю группу $H_k(S^k)$, так как $H_k(S^k) = G$ — одномерное линейное пространство. Итак, j_* — эпиморфизм, а потому $H_k(X) = G \oplus H_k(X \setminus D)$, что и завершает доказательство леммы.

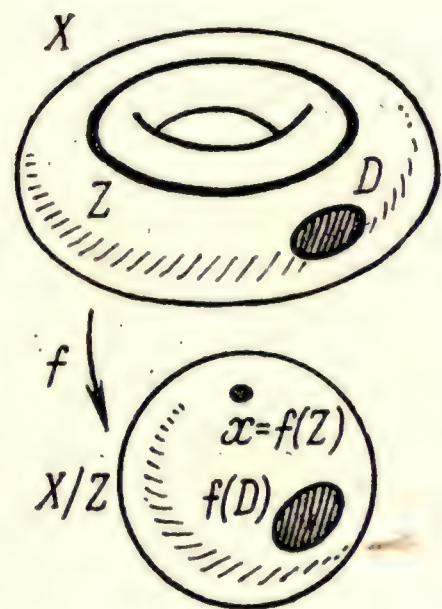


Рис. 52

Переходим к доказательству предложения 1. Рассмотрим поверхность X_0 и в качестве Z возьмем объединение множества $A \subset X_0$ и множества всех сингулярных точек минимальной поверхности X_0 . В силу теоремы существования минимальной пленки (см. [7, 8, 13—15]) мы знаем, что k -мерный объем множества Z равен нулю, следовательно, $H_k(Z) = 0$. Предположим, что $H_k(X_0) \neq 0$, и пусть l — любой ненулевой элемент этой группы. Приведем это предположение к противоречию. В силу леммы 1 существует диск D в X_0 такой, что $l \notin \text{Im } i_*$, где $i: X_0 \setminus D \rightarrow X_0$ — вложение, а i_* — индуцированный гомоморфизм $H_k(X_0 \setminus D) \rightarrow H_k(X_0)$. Рассмотрим коммутативную диаграмму:

$$\begin{array}{ccccccc}
 0 \rightarrow H_k(X_0 \setminus D) & \xrightarrow{i_*} & H_k(X_0) & \xrightarrow{j_*} & G & \xrightarrow{\varphi_*} & H_{k-1}(X_0 \setminus D) \rightarrow H_{k-1}(X_0) \rightarrow 0 \\
 & & & & \uparrow & & \uparrow \\
 & & & & \alpha_* & & \psi_* \\
 & & & & \leftarrow H_{k-1}(A) & \rightarrow & \beta_*
 \end{array}$$

Ясно, что $\text{Im } j_* \neq 0$, и так как $\dim G = 1$, то $\text{Im } j_* = G$, т. е. ψ_* — мономорфизм. Так как $\text{vol}_k X_0 > \text{vol}_k(X_0 \setminus D)$, то $X_0 \setminus D$ не является G -поверхностью с границей A , поскольку X_0 реализует абсолютный минимум k -мерного объема по всем поверхностям, заклеивающим границу A . Следовательно, существует элемент $\gamma \in H_{k-1}(A)$, $\gamma \neq 0$, такой, что $\alpha_* \gamma \neq 0$. С другой стороны, $\beta_* \gamma = 0$, так как X_0 является G -поверхностью с границей A . В силу ком-

мутативности имеем: $\psi_* \alpha_* \gamma = \beta_* \gamma = 0$, т. е. ψ_* имеет ненулевое ядро, т. е. элемент $\alpha_* \gamma$ отличен от нуля. С другой стороны, j_* — эпиморфизм, следовательно, ψ_* — мономорфизм. Утверждение доказано.

Перейдем к доказательству предложения 2. Пусть X_0 — глобально минимальная поверхность, реализующая при вложении $X_0 \rightarrow M$ нетривиальную группу P в $H_k(M)$. Предположим, что в $H_k(X_0)$ существует ненулевой элемент l , переходящий в ноль при гомоморфизме $i_* : H_k(X_0) \rightarrow H_k(M)$. В силу леммы 1 существует диск D в $X_0 \setminus Z$ (где Z — объединение границы A и множества сингулярных точек пленки X_0) такой, что при гомоморфизме $p_* : H_k(X_0 \setminus D) \rightarrow H_k(X_0)$, индуцированном вложением $p : X_0 \setminus D \rightarrow X_0$, элемент l не принадлежит образу гомоморфизма p_* и $\text{Codim Im } p_* = 1$ в $H_k(X_0)$. Тогда существует разложение группы $H_k(X_0)$ в прямую сумму векторных подпространств: $H_k(X_0) = G \oplus \text{Im } p_*$, где одномерное подпространство G натянуто на вектор l . Пусть α — произвольный ненулевой элемент из P , тогда $\alpha = i_* \varphi$ для некоторого φ , где $\varphi \in H_k(X_0)$. В силу разложения получаем: $\varphi = \lambda l + \varphi'$, где $\varphi' \notin \text{Im } p_*$ и $\varphi' \neq 0$, так как в противном случае $\alpha = i_*(\lambda l) = 0$ в силу исходного предположения $i_* l = 0$. Отсюда $\alpha = i_* \varphi = i_*(\lambda l + \varphi') = i_* \varphi' = i_* p_* \varphi''$, где $\varphi'' \in H_k(X_0 \setminus D)$. Итак, $P \subseteq \text{Im}(ip)_*$. Таким образом, поверхность $X_0 \setminus D$ при ее вложении ip в многообразие M реализует подгруппу P , т. е. принадлежит классу $[P]$ в обозначении предложения 2. Однако это противоречит условию абсолютной минимальности X_0 в классе $[P]$, поскольку $\text{vol}_k X_0 \setminus D < \text{vol}_k X_0$. Предложение 2 доказано.

Верно даже более сильное утверждение. Если рассмотреть границу A , фиксированную подгруппу P в $H_k(M)$, то можно рассмотреть класс $[A, P]$ всех G -поверхностей с границей A , реализующих, кроме того, подгруппу P . Оказывается, что предложение 2 верно и для случая глобально минимальных поверхностей из класса $[A, P]$. Чтобы доказать это, нужно дополнить приведенное выше доказательство. Мы доказали, что $X_0 \setminus D$ реализует P . Осталось доказать, что $X_0 \setminus D$ является G -поверхностью с границей A . Рассмотрим коммутативную диаграмму:

$$\begin{array}{ccccccccc}
 0 & \rightarrow & H_k(X_0 \setminus D) & \xrightarrow{p_*} & H_k(X_0) & \xrightarrow{\tau_*} & G & \xrightarrow{\xi_*} & H_{k-1}(X_0 \setminus D) & \xrightarrow{\rho_*} & H_{k-1}(X_0) & \rightarrow & 0 \\
 & & \downarrow \varphi_* & & \downarrow i_* & & & & \uparrow \psi_* & & \uparrow \omega_* & & \\
 & & & & H_k(M) & \leftarrow & & & H_{k-1}(A) & \rightarrow & & &
 \end{array}$$

Здесь $P \subseteq \text{Im } p_*$. Как и при доказательстве предложения 1, получаем, что ρ_* — мономорфизм, а тогда $\psi_* H_{k-1}(A) = 0$, так как $\omega_* = \rho_* \psi_*$ и $\omega_* H_{k-1}(A) = 0$. Таким образом, $X_0 \setminus D \in [A, P]$, что невозможно, так как $\text{vol}_k X_0 \setminus D < \text{vol}_k X_0$. Доказательство закончено.

Итак, глобально минимальные G -поверхности с границей A не содержат (в указанных предположениях) нетривиальных k -мерных циклов, т. е. k -мерных «пузырей», а глобально

минимальные поверхности, реализующие нетривиальные циклы в объемлющем многообразии, не содержат в себе никаких растворимых циклов, т. е. каждый пузырь минимальной пленки обязательно обволакивает какую-то нетривиальную дырку-пузырь в объемлющем многообразии. Хотя с интуитивной точки зрения эти утверждения естественны и трудно представить себе, что в некоторых ситуациях они могут быть неверными, тем не менее наша интуиция и здесь оказывается недостаточной, и в действительности существуют некоторые G -поверхности (для подходящих групп G), содержащие в себе нетривиальные k -мерные пузыри-циклы. Предъявим соответствующие примеры. В качестве группы G возьмем теперь не одномерную группу над полем, а группу U комплексных чисел с модулем, равным единице, т. е. $\{e^{i\varphi}\}$. Эта группа представима в виде $\mathbb{R}^1(\text{mod } 1)$ и гомеоморфна окружности. Рассмотрим класс минимальных U -поверхностей с границей A . Оказывается, существуют глобально минимальные U -поверхности с границей A , обладающие не только нетривиальной группой k -мерных гомологий (т. е. содержащие в себе пузыри-циклы размерности k), но и реализующие нетривиальные циклы в объемлющем многообразии, т. е. обволакивающие k -мерные дырки в M . В качестве M рассмотрим прямое произведение $M = S^1 \times \mathbb{R}P^{2s}$, где $\mathbb{R}P^{2s}$ — вещественное проективное пространство. В качестве границы A возьмем проективное подпространство $\mathbb{R}P^{2s-1}$, вложенное в $\mathbb{R}P^{2s}$ стандартным образом и являющееся его $(2s-1)$ -мерным остовом. В качестве U -поверхности X_0 возьмем пространство $\mathbb{R}P^{2s}$. Ясно, что $H_{2s-1}(X_0, U) = 0$, а группа $H_{2s-1}(A, U)$ нетривиальна и изоморфна группе U . Следовательно, при вложении $A \rightarrow X_0$ вся группа $H_{2s-1}(A, U)$ отображается в ноль, т. е. X_0 — U -поверхность с границей A . Поскольку X_0 является сомножителем в прямом произведении $S^1 \times \mathbb{R}P^{2s}$, то легко показать, что X_0 — глобально минимальная поверхность среди всех U -поверхностей с границей A . С другой стороны, $H_{2s}(X_0, U) = H_{2s}(\mathbb{R}P^{2s}, U) = \mathbb{Z}_2$, т. е. группа нетривиальна, и легко видеть, что X_0 реализует нетривиальную подгруппу $\mathbb{Z}_2$ в группе гомологий M .

5. Минимальные конусы, связанные с особыми точками минимальных поверхностей.

Вернемся к вопросу о физической реализации мыльных пленок, как содержащих, так и не содержащих в себе замкнутые пузыри-циклы. Этот вопрос тесно связан с поведением мыльных пленок вблизи их особых точек. Опираясь на принципы Платона, можно считать, что в достаточно малой окрестности каждой сингулярной точки минимальная поверхность состоит из нескольких гладких кусков двумерной поверхности, которые можно считать приблизительно плоскими. Как мы увидим далее, это свойство может не выполняться для минимальных поверхностей размерностей больших, чем два. Существуют «многомерные» сингулярные точки, никакая сколь угодно малая

окрестность которых не состоит из плоских кусков. Возвратимся к двумерному случаю. Рассмотрим сферу малого радиуса с центром в особой точке и изучим ее пересечение с поверхностью. Можно считать, что пересечение состоит из конечного числа гладких дуг, содержащихся в сфере и образующих на ней некоторую сеть.

Легко убедиться, что каждый отрезок такой дуги должен быть отрезком некоторого экватора, т. е. окружности, получающейся в результате пересечения сферы плоскостью, прохо-

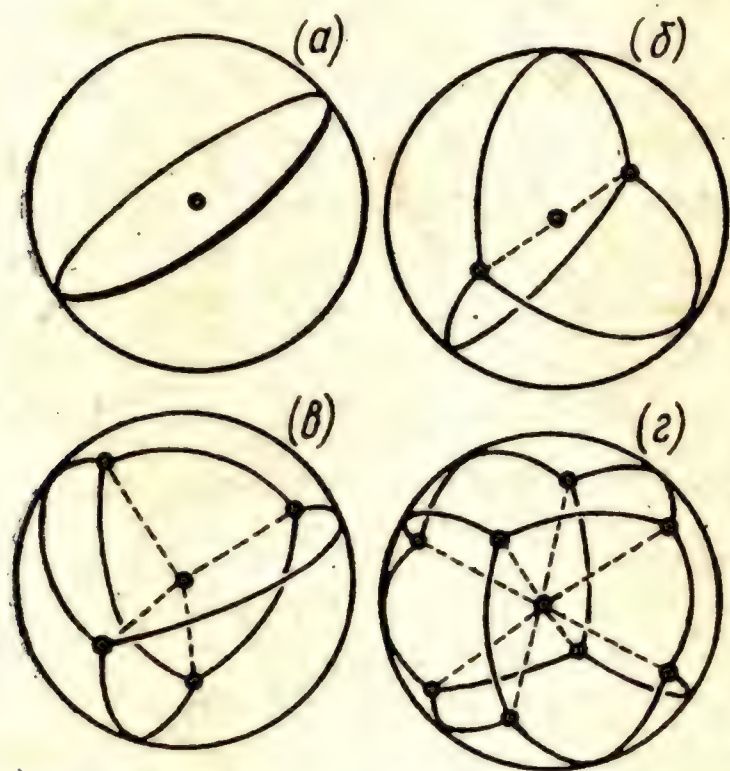


Рис. 53 (а) — (г)

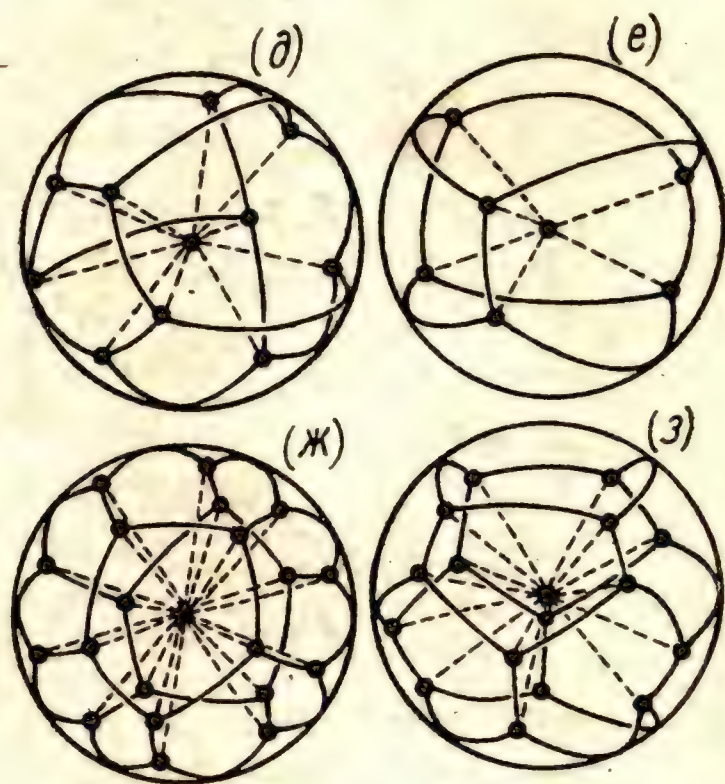


Рис. 53 (д) — (з)

дящей через ее центр. В самом деле, так как мы считаем, что пленка около особой точки состоит из плоских кусков, то они сходятся в особой точке и вся часть пленки внутри сферы является конусом с вершиной в особой точке. Конус имеет своим «основанием» систему гладких дуг, образующих сеть на сфере. Мыльная пленка, заключенная внутри сферы, образована всевозможными радиусами, исходящими из центра сферы и кончающимися на одномерной сети. Некоторые радиусы могут быть сингулярными ребрами пленки, где сходится несколько ее листов. Пленка составлена из нескольких секторов, частей обычного плоского диска. Имеется простая связь между длиной l одномерной сети и площадью S пленки, заключенной в сфере. Ясно, что $S = \frac{1}{2} r_0 l$, где r_0 — радиус сферы. Отсюда следует, что сеть состоит из отрезков экваторов, т. е. геодезических на сфере. Если допустить, что часть дуги не является геодезической, то имеется малая деформация, не смещающая ее концов, но уменьшающая длину. Поэтому существует сокращающая деформация, уменьшающая и площадь двумерной пленки-конуса, так как уменьшение l вызывает уменьшение S . Таким образом, если в какой-то точке на сфере сходится несколько гладких дуг сети, то число их равно трем и они образуют друг с другом одинаковые углы, равные $2\pi/3$.

Поставим вопрос: каковы те одномерные сети, которые получаются как результат пересечения мыльной пленки с малой сферой, центр которой помещен в какой-то сингулярной точке пленки? Ответив на вопрос, мы получим описание локальной структуры пленки в окрестности ее особой точки, не лежащей на граничном контуре, затягиваемом пленкой. Обнаруженные нами факты (а именно: 1) каждая гладкая дуга сети является частью экватора; 2) в каждой особой точке сети, в ее узле, могут сходиться только 3 дуги, образующие углы $2\pi/3$) позволяют дать полный список всех таких возможных конфигураций — сетей на сфере. Таких конфигураций есть ровно 10, и все они приведены на рис. 53. Чтобы убедиться в отсутствии других сетей, достаточно воспользоваться элементарными соображениями сферической тригонометрии.

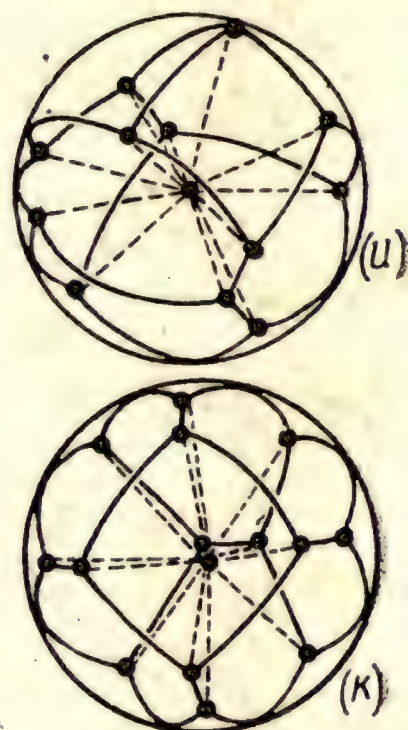


Рис. 53(и), (к)

Теперь выясним: действительно ли все эти конфигурации реализуются как пересечения некоторых мыльных пленок с малой сферой, центрированной в особой точке? Оказывается, большинство из указанных конфигураций таковыми не являются. Конечно, если мы рассмотрим прямолинейные конусы над этими сетями с вершиной в центре сферы, то они имеют нулевую среднюю кривизну во всех своих регулярных точках. Конусы показаны на рис. 53. Если пренебречь множеством меры нуль, то они являются минимальными поверхностями. Однако мы уже хорошо знаем, что игнорировать особые точки пленки нельзя и что условие минимальности накладывает жесткие ограничения на структуру сингулярностей мыльных пленок, особенно, если мы интересуемся устойчивыми пленками.

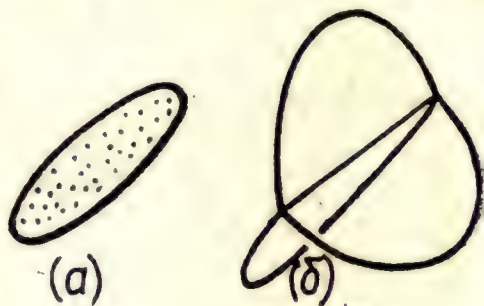


Рис. 54(а), (б)

Основной интерес для нас представляют пленки, минимальные в следующем смысле: любое малое возмущение с малым носителем не уменьшает площадь пленки. С точки зрения этого определения большинство конусов на рис. 53 минимальными не являются. Во-первых, ясно, что первые три поверхности, затягивающие соответственно плоскую окружность, три дуги, сходящиеся под равными углами в двух вершинах, правильный тетраэдр, являются устойчивыми минимальными пленками. Во-вторых, оказывается, все остальные конусы не минимальны. Для каждого из них есть сокращающая деформация, возмущение, уменьшающее площадь. При этом вершина конуса распадается, «раздувается» в топологически более сложное образование, но зато имеющее особые точки, удовлетворяющие принципам Платона. На рис. 54 изображены те реальные мыль-

ные пленки, которые получаются при попытке реализовать минимальные поверхности с границей, показанной на рис. 53 (см. [11]). Структура этих поверхностей достаточно сложна,

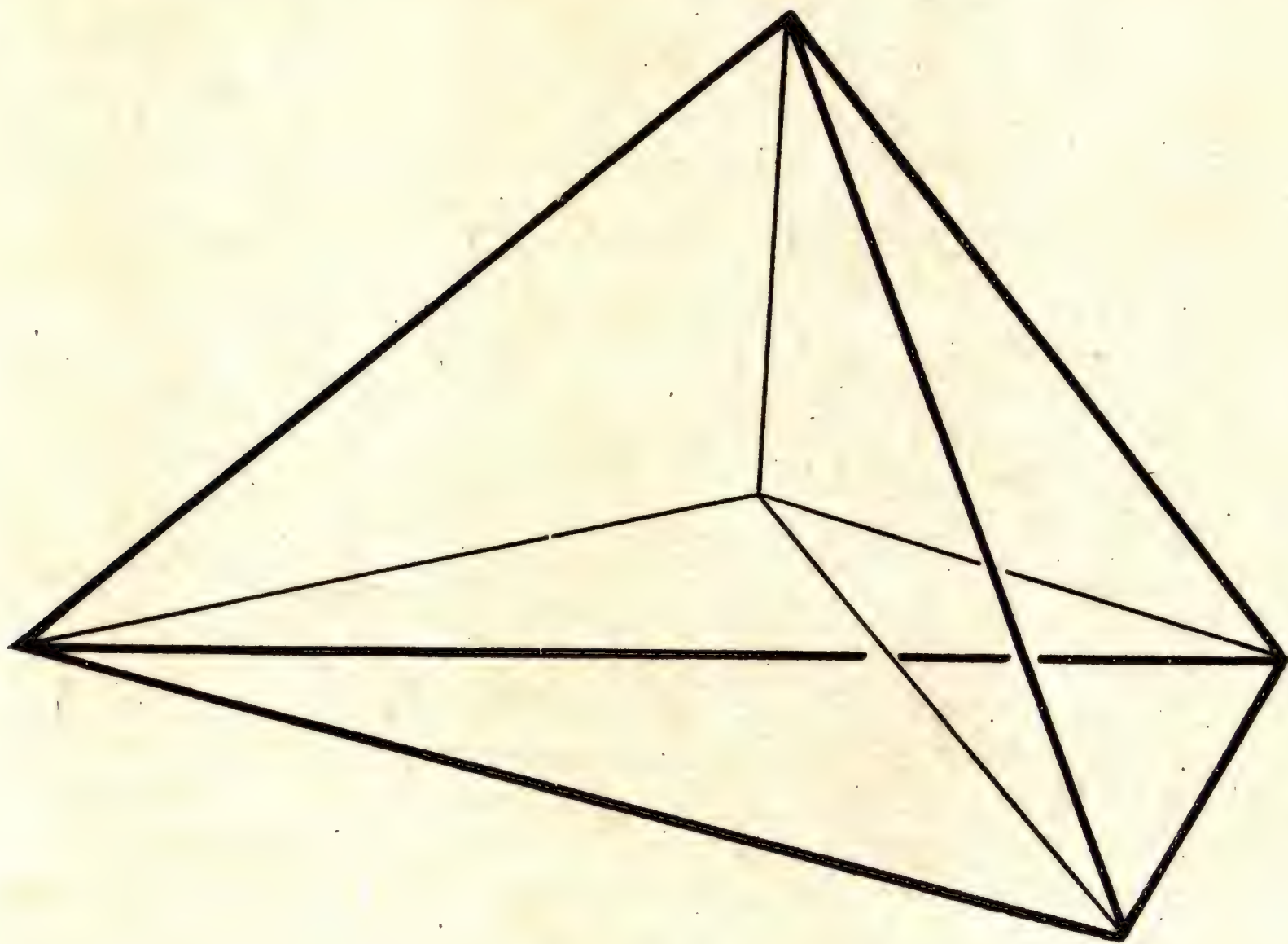


Рис. 54(в)

однако все их особые точки либо расположены на трехкратных сингулярных ребрах, либо являются четырехкратными сингулярностями, в которых сходятся под равными углами 4 ребра

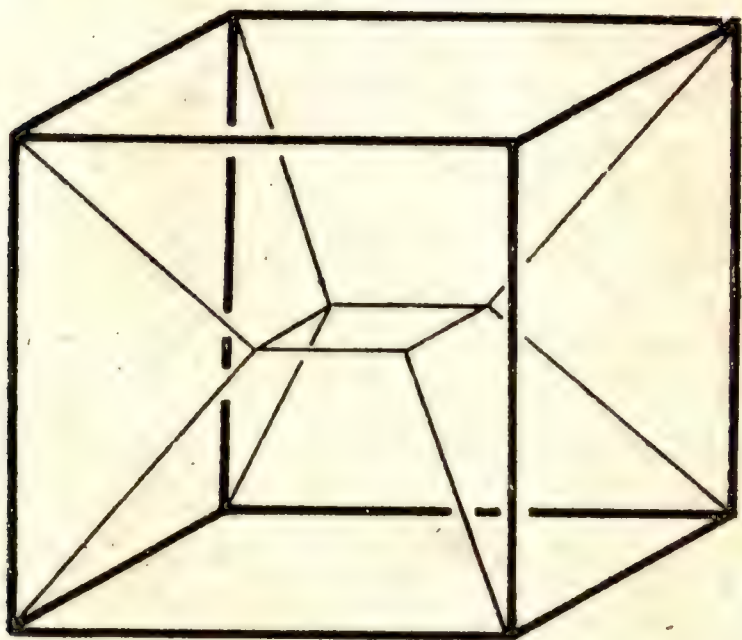


Рис. 54(г)

пленки. Заинтересованный читатель может, проявив достаточное терпение и аккуратность, получить реальные мыльные пленки описанного топологического типа, используя проволоочные контуры, показанные на рисунке. Эти пленки устойчивы, и любое малое возмущение не уменьшает их площадь. Пленки, показанные на рис. 54(г)—54(к), уже не могут рассматриваться как малые, поскольку они, как выяснилось, не моделируют структуру бесконечно малой окрестности какой-

либо сингулярной точки на мыльной пленке. В отличие от них первые три пленки (а), (б), (в) допускают преобразование подобия (сжатие с центром в начале координат), переводящее их

в себя и уменьшающее их размер. Эти пленки моделируют структуру пленки в бесконечно малой окрестности сингулярности. Здесь мы коснулись интересного вопроса о минимальных

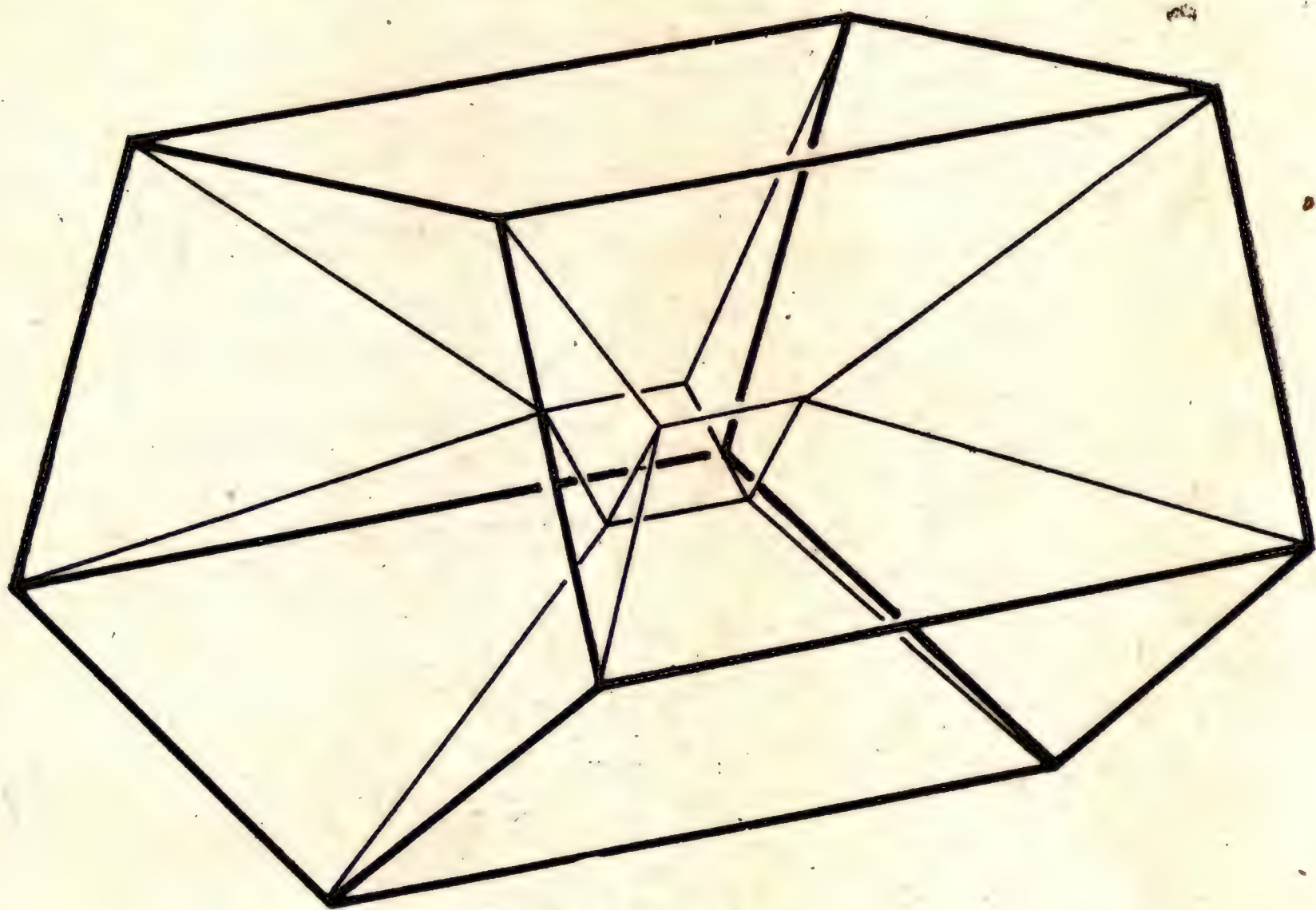


Рис. 54(д)

конусах в $\mathbb{R}^n$, с вершиной в начале координат, имеющих «основанием» — границей — минимальное подмногообразие A^{n-2} в сфере S^{n-1} .

6. Многомерные минимальные конусы.

Рассмотрим в $\mathbb{R}^n$ сферу S^{n-1} радиуса R с центром в начале координат и пусть A^{n-2} — $(n-2)$ -мерное компактное подмногообразие в сфере, являющееся почти всюду (за исключением множества точек меры нуль) гладким $(n-2)$ -мерным гладким подмногообразием в сфере. Рассмотрим конус CA с вершиной в точке O и с основанием A . Конус образован радиусами, идущими из центра в точки множества A . Фиксируем границу A конуса CA и предположим, что конус является минимальной поверхностью в том смысле, что любое его возмущение, оставляющее на месте границу, увеличивает (не уменьшает) объем, площадь в двумерном случае. Возмущение конуса будем понимать в смысле пунктов 1, 2 настоящего параграфа. Ясно, что далеко не при всяком выборе границы A конус CA будет минимален.

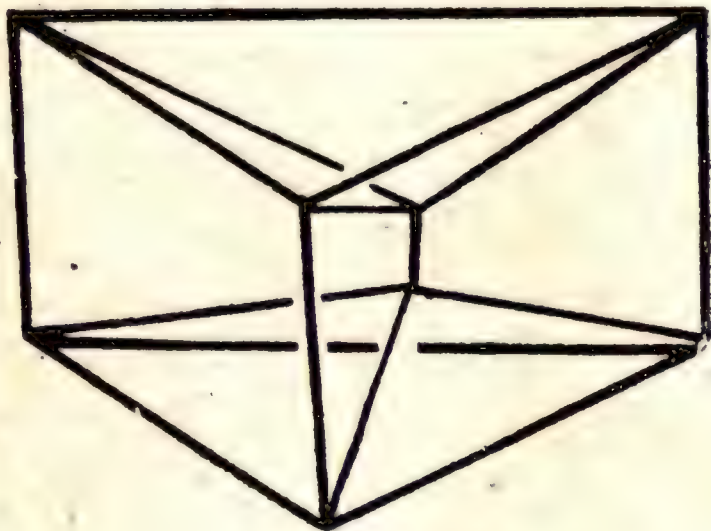


Рис. 54(е)

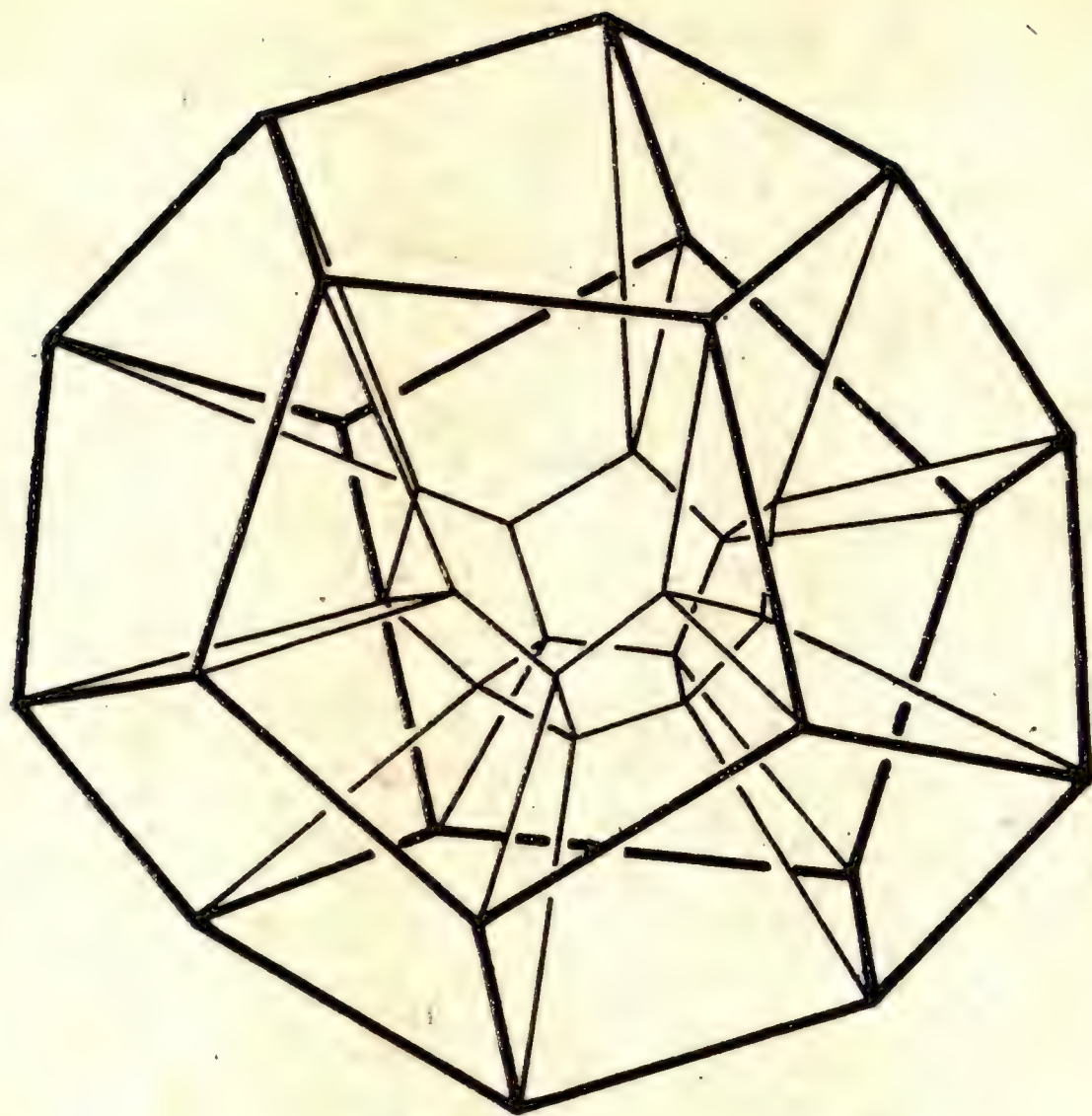


Рис. 54(ж)

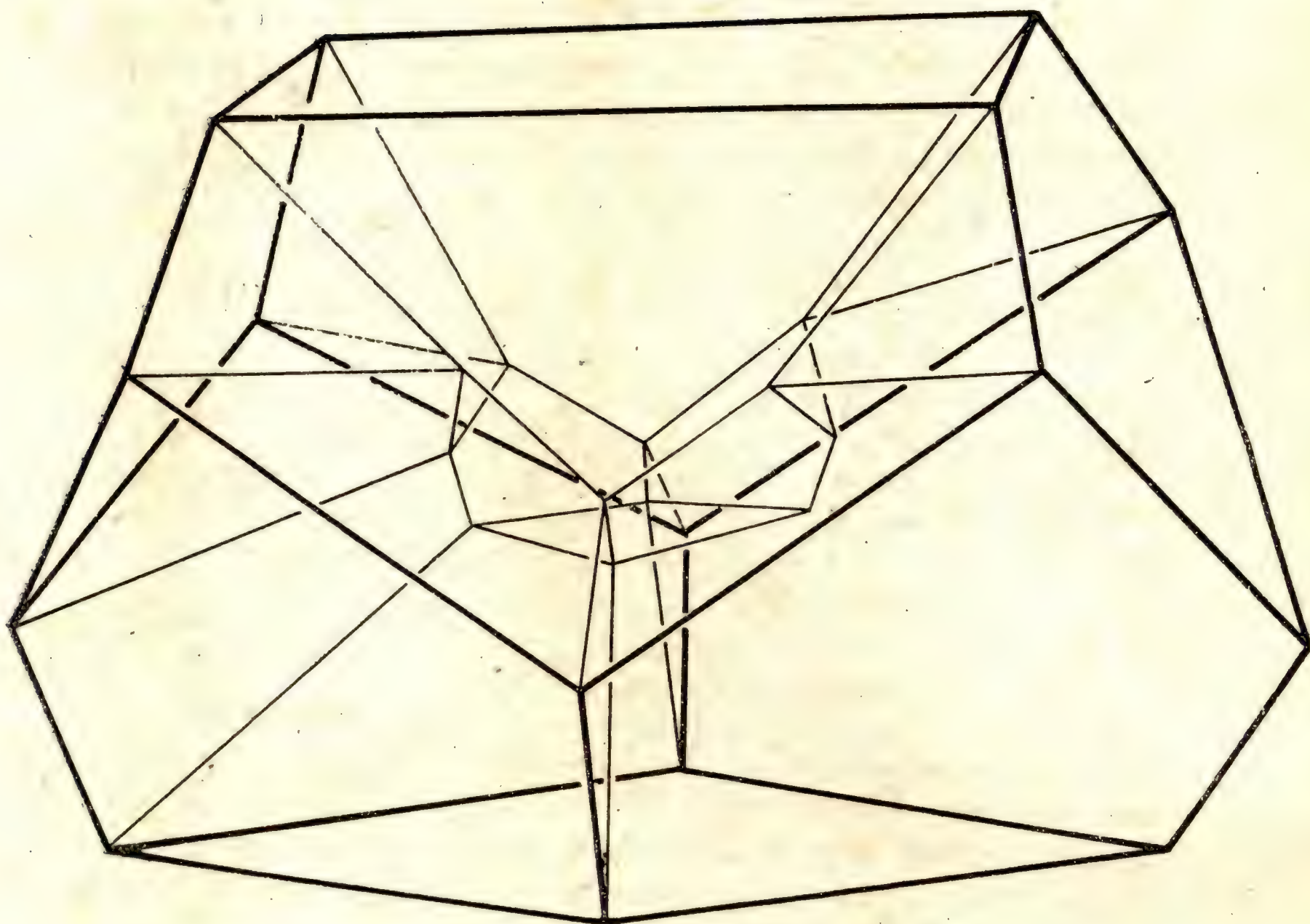


Рис. 54(з)

Лемма 1. Если конус CA является минимальной поверхностью в $\mathbb{R}^n$, то его граница A является минимальной поверхностью в сфере S^{n-1} во всех своих регулярных точках.

В частности, за исключением множества точек меры нуль, граница A является локально минимальным подмногообразием

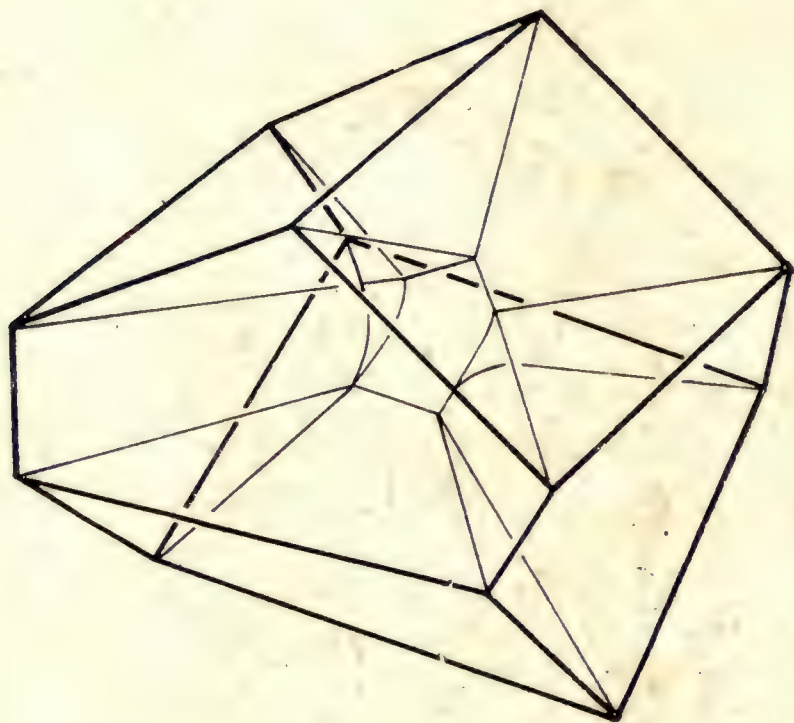


Рис. 54(и)

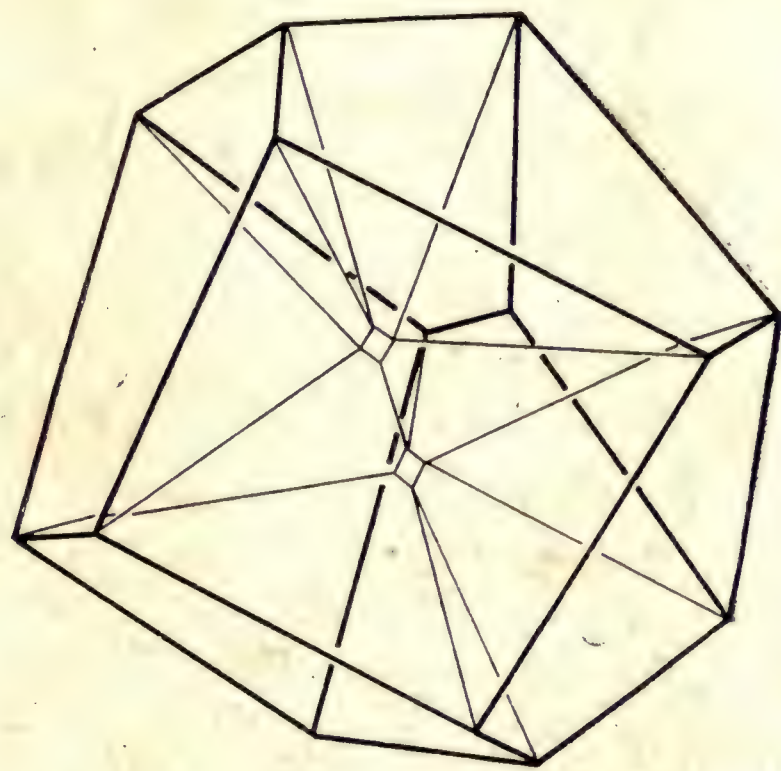


Рис. 54(к)

в сфере, т. е. его средняя кривизна равна нулю. Из элементарных свойств интеграла следует, что $(n-1)$ -мерный объем конуса CA связан с $(n-2)$ -мерным объемом его границы A следующим соотношением: $\text{vol}_{n-1} CA = \frac{R}{n-1} \text{vol}_{n-2} A$, где R —

радиус сферы. Следовательно, любое малое возмущение границы, уменьшающее ее объем, индуцирует (по радиусам) аналогичное возмущение всего конуса, уменьшающее его объем. Итак, если допустить существование сокращающей деформации для границы, то сокращающая деформация найдется и для конуса, что противоречит его предполагаемой минимальности и доказывает лемму.

Отсюда следует уже знакомое нам утверждение о том, что пересечение двумерной мыльной пленки со сферой бесконечно малого радиуса с центром в особой точке пленки можно считать состоящим из отрезков экваторов, сходящихся в трехкратных точках под равными углами.

Для простоты ограничимся случаем, когда граница конуса является гладким замкнутым подмногообразием в сфере. Вопрос: существуют ли в $\mathbb{R}^n$ конусы с вершиной в начале координат и являющиеся минимальными в том смысле, что любое малое их возмущение увеличивает объем? Простейшим таким конусом является стандартный плоский $(n-1)$ -мерный евклидов диск, проходящий через начало координат и пересекающий сферу по экватору. Поэтому уточним вопрос: существуют ли нетривиальные минимальные конусы, отличные от стандартного диска? Ответ является настолько неожиданным, что мы

остановимся на этом подробнее. Ясно, что в $\mathbf{R}^3$ никаких минимальных конусов (кроме плоского диска), имеющих в качестве границы гладкое одномерное подмногообразие, не существует. Это следует из того, что граница должна быть траекторией локально-минимальной длины, а все геодезические на сфере исчерпываются экваторами. Следовательно, конус является плоским диском. С повышением размерности в сфере S^{n-1} появляются локально минимальные подмногообразия A^{n-2} , не совпадающие с экваториальными гиперсферами. Так, например, в S^3 кроме обычных экваторов содержится локально минимальный двумерный тор, вложенный следующим образом. Пусть $S^3 \subset \mathbf{R}^4 = \mathbf{C}^2(z, w)$, тогда тор T^2 можно задать как пересечение сферы $S^3 = \{|z|^2 + |w|^2 = 1\}$ с поверхностью $M^3 = \{|z| = |w|\}$. Сфера является объединением двух полноторий: $|z| \leq |w|$ и $|z| \geq |w|$. Тор является их общей границей. Если φ и ψ — угловые координаты, то точки тора задаются так: $(z, w) = (e^{i\varphi}, e^{i\psi}) / \sqrt{2}$. Метрика, индуцированная на торе, является плоской евклидовой метрикой, так как $dzd\bar{z} + dw\bar{w} = (d\varphi^2 + d\psi^2)/2$. Легко проверить, что тор является локально минимальной поверхностью в сфере. Ниже мы вернемся к

этому вопросу. Однако можно убедиться также в том, что минимальная поверхность, расположенная внутри сферы и имеющая тор своей границей, не является конусом. Оказывается, в $\mathbf{R}^4$, как и в $\mathbf{R}^3$, нет нетривиальных конусов с центром в начале координат. С увеличением размерности картина меняется. Оказывается, если размерность велика, то существуют минимальные нетривиальные конусы [22]. Чтобы получить интуитивное представление о причинах этого любопытного эффекта, рассмотрим, для определенности, границу A^{n-2} фиксированного топологического типа. Простейшими многообразиями, следующими за сферами и допускающими локально минимальные вложения в сферу, являются произведения

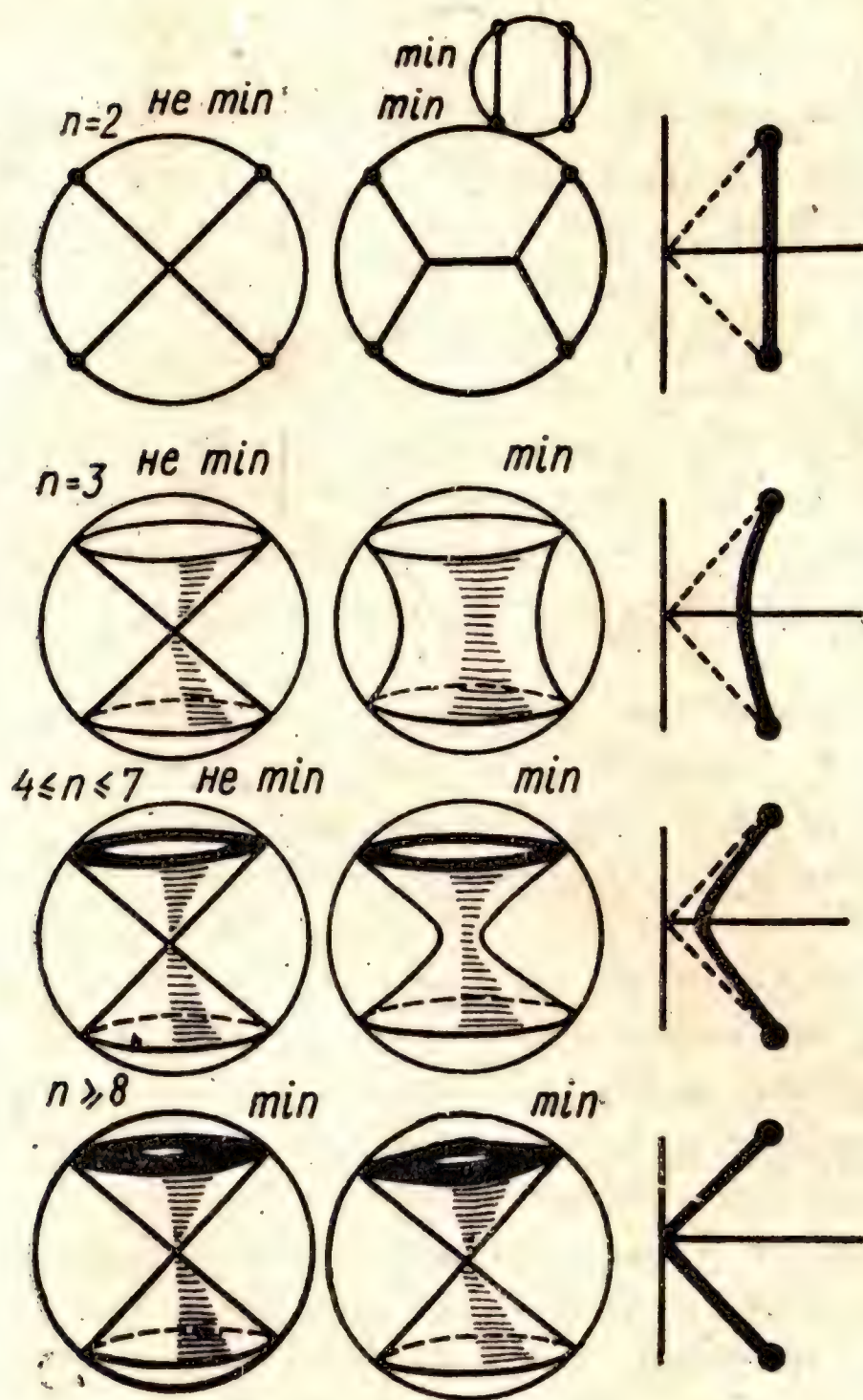


Рис. 55

сфер $S^p \times S^q$, где $p+q=n-2$. Пусть $n=2$, рассмотрим окружность S^1 и $A=S^0 \times S^0 \subset S^1$ (рис. 55). В этом случае конус с границей A является объединением двух диаметров. Эта одномерная «поверхность» не минимальна, четырехкратная точка распадается в объединение двух трехкратных. Реальная минимальная траектория с границей A является объединением двух параллельных отрезков — «одномерным цилиндром» (рис. 55). Пусть теперь $n=3$. В качестве A в сфере S^2 возьмем $S^1 \times S^0$. Ясно, что двумерный конус с границей A снова не минимален, так как существует сокращающая деформация в окрестности вершины, уменьшающая площадь. В результате этой вариации особая точка — вершина конуса — распадается и превращается в окружность, являющуюся горловиной катеноида — реальной минимальной пленки с границей $S^1 \times S^0$. Сравнивая двумерную минимальную пленку — катеноид — с одномерной, т. е. с двумя параллельными отрезками, мы замечаем, что двумерная пленка провисает, прогибается по направлению к началу координат в отличие от одномерной пленки. Оказывается, это провисание минимальной пленки по направлению к началу координат увеличивается с ростом размерности. Рассмотрим S^3 и в ней тор T^2 , описанный выше, в качестве границы. Можно подсчитать, что трехмерная минимальная пленка (рис. 55) имеет еще более узкую горловину, чем двумерный катеноид. На рис. 55 в первом столбце условно изображены конусы с границей $S^p \times S^q$, во втором столбце — реальные минимальные поверхности с этой же границей, в третьем — плоские сечения (образующие) этих минимальных поверхностей двумерными полуплоскостями, проходящими через ось симметрии поверхности. Сечение имеет вид траектории, которая с ростом размерности провисает все больше и больше по направлению к точке O . Следовательно, в этом монотонном процессе наступает момент, когда минимальная поверхность провиснет настолько, что ее горловина схлопнется, стянется в точку и пленка превратится в конус. Происходит это начиная с размерности 8. До размерности 7 включительно не существует минимальных конусов, отличных от плоского диска (Альмгрен, Саймонс).

Теорема 1 (см. [23]). Пусть A^{n-2} — замкнутое гладкое локально-минимальное подмногообразие в сфере S^{n-1} , стандартно вложенной в $\mathbb{R}^n$. Пусть A не является вполне геодезической сферой, т. е. экватором. Тогда при $n \leq 7$ конус CA (т. е. $(n-1)$ -мерная поверхность, образованная всеми радиусами, идущими из начала координат в точки многообразия A) не является минимальным, т. е. существует сокращающая деформация конуса, уменьшающая его объем.

Это утверждение носит общий характер в том смысле, что оно справедливо не только для евклидова пространства, но и для односвязных полных римановых многообразий M^n постоянной секционной кривизны (сфера и пространство Лобачевского). А именно, пусть S_R^{n-1} — геодезическая сфера в M^n

радиуса R , образованная концами геодезических радиусов длины R , исходящих из фиксированной точки — центра сферы. Пусть CA — конус над A , образованный геодезическими, идущими из центра в точки гладкого замкнутого локально минимального подмногообразия A^{n-2} , вложенного в сферу S_R^{n-1} . Тогда, если A не является вполне геодезической сферой, то конус CA не минимален при $n \leq 7$. Более того, это утверждение верно для любых римановых пространств, полученных из $\mathbb{R}^n$ домножением евклидовой метрики на положительную функцию $f(R)$, зависящую только от радиуса R (доказано студенткой МГУ Хорьковой).

Вернемся к случаю $\mathbb{R}^n$. Начиная с $n=8$ в $\mathbb{R}^n$ уже существуют конусы, являющиеся не только локально, но и глобально минимальными. Предъявим пример. В качестве границы рассмотрим произведение сфер $S^3 \times S^3$ одинакового радиуса, естественно вложенное в S^7 как локально-минимальное подмногообразие [23, 24]. Укажем это вложение. Рассмотрим в $\mathbb{R}^{2p}$ сферу S^{2p-1} единичного радиуса: $x_1^2 + \dots + x_{2p}^2 = 1$ и пересечем ее с конусом $x_1^2 + \dots + x_p^2 = x_{p+1}^2 + \dots + x_{2p}^2$. Это пересечение диффеоморфно $S^{p-1} \times S^{p-1}$. Вычислением убеждаемся, что подмногообразие $S^{p-1} \times S^{p-1}$ локально-минимально в сфере. Ниже мы еще раз убедимся в этом на основе других соображений.

7. Минимальные поверхности, инвариантные относительно действия групп Ли.

Рассмотрим гладкое риманово многообразие M , на котором гладко действует его группа изометрий. Путь I_0 — связная компонента единицы этой группы Ли и пусть G — некоторая связная компактная подгруппа в группе I_0 . Пусть $X \subset M$ — поверхность, инвариантная относительно действия группы G . Такие поверхности называются G -инвариантными, они расслаиваются на орбиты действия группы G . Рассмотрим многомерную задачу Плато. Пусть в M фиксирован «контур» — замкнутое подмногообразие A^{n-1} , инвариантное относительно действия G . Естественно ожидать, что если граница обладает некоторой группой симметрий (G -инвариантна), то такую же группу симметрий должна иметь и минимальная поверхность, затягивающая границу. Эквивариантная задача Плато: в каких случаях можно гарантировать существование G -инвариантной минимальной поверхности X^{n-1} , затягивающей (в указанных выше смыслах) G -инвариантную границу.

Пользуясь средствами дифференциальной геометрии, можно доказать (см., например, [25]), что: G -инвариантное подмногообразие X в многообразии M локально-минимально относительно всех достаточно малых вариаций в том и только в том случае, когда оно является локально-минимальным лишь по отношению ко всем достаточно малым эквивариантным вариациям (т. е. инвариантным относительно действия той же груп-

пы). Итак, для проверки локальной минимальности (в смысле равенства нулю средней кривизны) G -инвариантной поверхности достаточно убедиться, что она минимальна относительно более узкого класса вариаций, а именно относительно G -инвариантных возмущений. Это позволяет, как мы увидим, редуцировать задачу нахождения G -инвариантных минимальных поверхностей в M и задаче нахождения минимальных поверхностей в пространстве орбит M/G .

Рассмотрим разбиение многообразия M в объединение орбит $G(x)$ действия группы G , где $x \in M$. Из разных точек «вырастают» орбиты, вообще говоря, разных размерностей, однако в множестве всех орбит естественно выделен некоторый подкласс орбит, называемых орбитами общего положения, или главными орбитами. Опишем этот класс. Напомним, что каждая орбита $G(x)$ допускает представление в виде однородного пространства $G/H(x)$, где $H(x)$ — стационарная подгруппа точки x , т. е. совокупность всех преобразований из G , оставляющих точку на месте. Множество классов смежности в группе G по подгруппе $H(x)$ и отождествляется с пространством $G/H(x)$. Оказывается, в M можно выделить такое максимальное открытое всюду плотное подмножество $\tilde{M}$, являющееся в действительности открытым гладким подмногообразием в M , что для любых двух точек x, y из этого подмножества соответствующие им стационарные подгруппы $H(x)$ и $H(y)$ сопряжены в G , т. е. существует элемент $g \in G$ такой, что $H(y) = gH(x)g^{-1}$. Орбиты $G(x)$, соответствующие точкам x из этого всюду плотного подмножества $\tilde{M}$, и называются главными, или орбитами общего положения. В частности, все эти орбиты попарно диффеоморфны и имеют одинаковую размерность, являющуюся максимальной в классе всех орбит. Все остальные орбиты назовем особыми. Множество особых орбит имеет меру нуль в пространстве всех орбит, и соответствующие им точки x из M заполняют замкнутое подмножество, n -мерный объем (мера) которого равен нулю. Здесь $n = \dim M$. Группа G предполагается компактной. Могут существовать особые орбиты, имеющие максимально возможную размерность, но не являющиеся главными. Однако поскольку они заполняют подмножество меры нуль, то в дальнейшем такие орбиты нас не будут интересовать.

Пусть X^q — произвольная G -инвариантная поверхность в M . Положим $k = q - s$, где s — размерность главной орбиты, т. е. k — коразмерность орбит общего положения в поверхности X . Так как каждая главная орбита $a = G(x)$ является гладким s -мерным подмногообразием в M , то можно вычислить ее s -мерный объем, который обозначим $v(a)$. Через a обозначена орбита $G(x)$, проходящая через точку x . Рассмотрим пространство орбит M/G . Вообще говоря, оно не является многообразием и может содержать особенности. Однако оно содержит открытое всюду плотное подмножество $\tilde{M}/G$, являющееся многообра-

зием. Пусть $\pi: M \rightarrow M/G$ — проекция M на M/G , сопоставляющая каждой точке x проходящую через нее орбиту $G(x) = a$. Ограничение π на открытое подмногообразие $\tilde{M}$ в M определяет гладкое расслоение со слоем $\pi^{-1}(a) = G(x)$ над базой $\tilde{M}/G$ (см. гл. 1, § 3). Хотя все слои $\pi^{-1}(a)$ над базой $\tilde{M}/G$ диффеоморфны друг другу, однако они имеют, вообще говоря, различные s -мерные объемы, если слои рассматривать

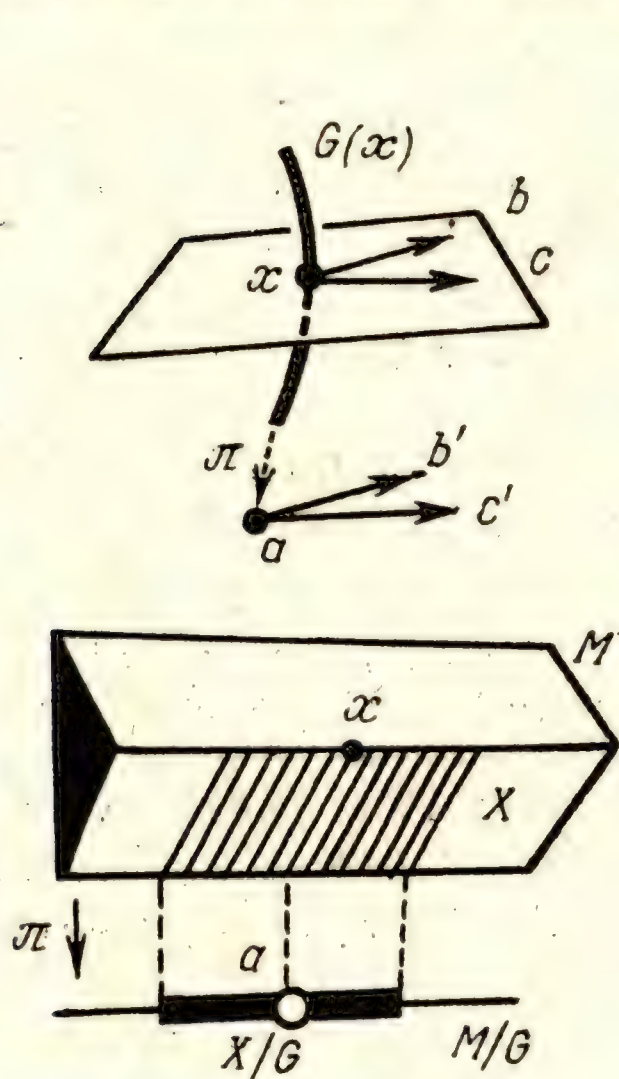


Рис. 56(а)

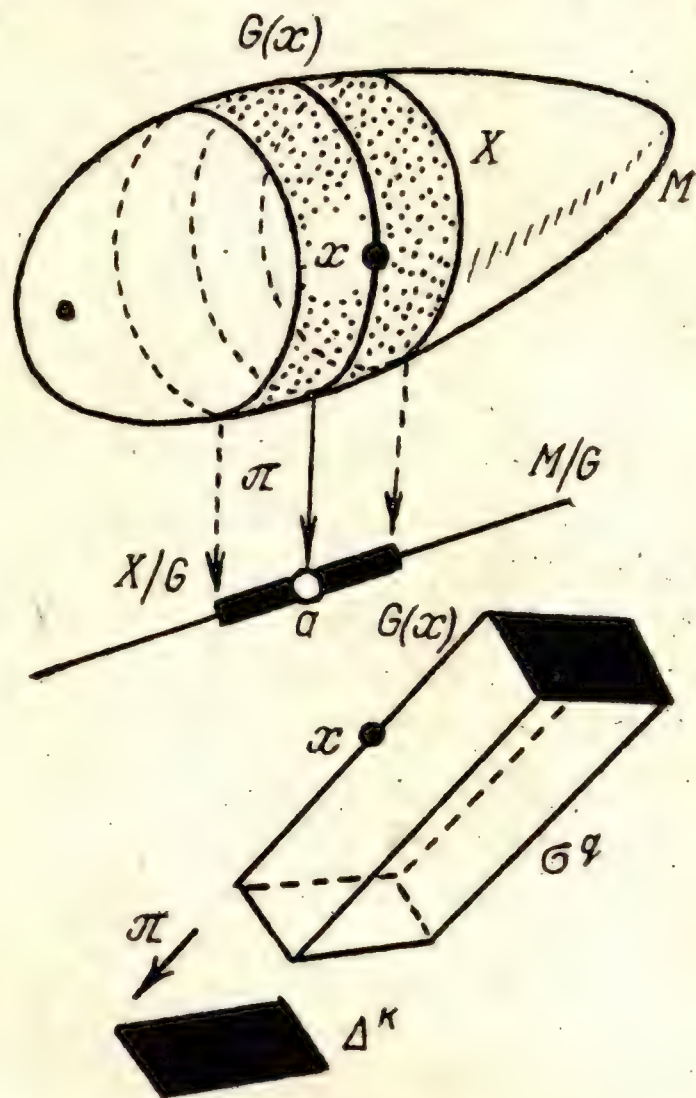


Рис. 56(б)

как подмногообразия $G(x)$ в M . Следовательно, сопоставляя каждой орбите ее объем, мы получаем гладкую функцию $v(a)$, определенную на $\tilde{M}/G$ (рис. 56). Она играет в дальнейшем важную роль [25, 26].

Определим полезную для наших целей риманову метрику на $\tilde{M}/G$. Используем риманову метрику ds , заданную на M , и функцию объема орбит $v(a)$. Пусть $\langle b, c \rangle$ — скалярное произведение касательных векторов b и c , определяемое метрикой на M . Фиксируем множество $(n-s)$ -мерных плоскостей, лежащих в касательных плоскостях к M и ортогональных к орбитам $G(x)$, т. е. к слоям расслоения $\pi: \tilde{M} \rightarrow \tilde{M}/G$. Пусть b', c' — два вектора, касательных к $\tilde{M}/G$ в точке a . Ясно, что в каждой точке $x = \pi^{-1}(a)$ однозначно определены их прообразы — два вектора b и c , принадлежащие к плоскости, ортогональной орбите $G(x)$, и проектирующиеся при отображении $d\pi_x$ в векторы b' и c' (рис. 56). Зададим невырожденное симметричное скалярное произведение $\langle, \rangle'$ векторов b' и c' , положив $\langle b', c' \rangle' = \langle b, c \rangle$. Мы определили риманову метрику на $\tilde{M}/G$, являющуюся проекцией исходной метрики ds . Обозначим

полученную метрику на $\tilde{M}/G$ через $d\tilde{s}$. Наконец, определим метрику dl на $\tilde{M}/G$, положив $dl = v(a)^{1/k} d\tilde{s}$ в каждой точке a из $\tilde{M}/G$. Так как $v(a)$ — положительная гладкая функция на $\tilde{M}/G$, то dl — невырожденная положительно-определенная метрика на фактор-многообразии $\tilde{M}/G$. В действительности, можно продолжить $v(a)$ до непрерывной функции на всем $\tilde{M}/G$, доопределив ее во всех точках a , являющихся особыми орбитами. Однако этим обстоятельством мы не будем пользоваться.

Выбор метрики dl обусловлен простым, но важным утверждением. Пусть X — компактное G -инвариантное подмногообразие в M ; рассмотрим фактор-пространство X/G , вложенное в M/G . Будем считать, что X находится в общем положении относительно множества особых орбит в M , и рассмотрим в X подмножество $\tilde{X} = X \cap \tilde{M}$, т. е. подмножество тех точек, через которые проходят главные орбиты. Так как X является G -инвариантной поверхностью, то $\tilde{X}$ представляется в виде объединения орбит; будем считать, что почти все эти орбиты являются главными, т. е. что $\tilde{X}$ — открытое всюду плотное подмногообразие в X . В частности, $\text{vol}_q X = \text{vol}_q \tilde{X}$. Напомним, что k — это коразмерность главной орбиты в X , следовательно, k — это размерность фактор-поверхности X/G .

Лемма 2. Если dl — метрика, введенная выше на $\tilde{M}/G$, то q -мерный объем G -инвариантной поверхности X^q в M относительно метрики ds в точности равен k -мерному объему фактор-поверхности $\tilde{X}/G$ в фактор-пространстве $\tilde{M}/G$ относительно метрики dl , где $k = \dim X - \dim G(x)$, $G(x)$ — главная орбита.

Доказательство вытекает из простых геометрических соображений. В самом деле, k -мерный объем (относительно dl) элементарного бесконечно малого куба Δ^k с ребром dl , расположенного в $\tilde{X}/G$, равен $(dl)^k$, где через dl условно обозначена также длина ребра куба. Отсюда $\text{vol}_k \Delta^k = v(a) (d\tilde{s})^k$. Из рис. 56(б) следует, что эта величина равна произведению объема основания параллелепипеда σ^q на объем его «образующей». При этом основанием параллелепипеда является элементарная площадка размерности k , объем которой, очевидно, равен $(d\tilde{s})^k$. Образующей σ^q является орбита действия группы. При этом объем орбиты $G(x)$ как раз равен $v(a)$, где $a = \pi(x)$.

Отсюда непосредственно вытекает следующее утверждение:

Предложение 1. Пусть X^q — некоторое G -инвариантное минимальное подмногообразие в M . Это подмногообразие локально-минимально (т. е. его средняя кривизна равна нулю) в M относительно метрики ds в том и только в том случае, когда фактор-многообразие $\tilde{X}/G$ локально-минимально в фактор-многообразии $\tilde{M}/G$ относительно метрики $dl = v^{1/k} \cdot d\tilde{s}$, где $k = \dim X - \dim G(x)$, т. е. число k равно коразмерности орбиты общего положения в поверхности X .

Уже при $k=0$ мы получаем интересные следствия. Если подмногообразие X^q считать для простоты связным, то равенство числа k нулю означает, что X является некоторой орбитой дей-

ствия группы G на объемлющем многообразии M . Возвращаясь к доказательству леммы 2, мы видим, что параллелепипед σ^1 вырождается в данном случае в свою образующую, т. е. в орбиту $G(x)$, совпадающую с X . Следовательно, локальная минимальность X в M относительно метрики ds эквивалентна условию, что функция объема v достигает экстремума в точке $a = \pi(X) = \pi G(x)$. Итак, все локально-минимальные главные орбиты $G(x)$ в M находятся так: нужно рассмотреть пространство главных орбит $\tilde{M}/G$, вычислить на нем функцию объема орбит $v(a)$ и найти все ее критические точки, т. е. такие, в которых $\text{grad } v(a) = 0$. Они и являются орбитами, являющимися локально-минимальными подмногообразиями в M . Простейшим примером является действие одномерной группы $G = \text{SO}(2)$ как группы вращений стандартной сферы S^2 вокруг вертикальной оси. Пространство орбит S^2/G здесь одномерно, гомеоморфно отрезку, пространство главных орбит гомеоморфно открытому интервалу. Функция объема определена на открытом интервале, является квадратным корнем из квадратичной функции и имеет ровно один экстремум — максимум, достигаемый в центре интервала. Максимум соответствует максимальной длине окружности. Соответствующая орбита является окружностью наибольшего радиуса — экватором. Следовательно, экватор — геодезическая в сфере.

При $k=1$ G -инвариантное подмногообразие X локально-минимально в M относительно метрики ds тогда и только тогда, когда траектория $\tilde{X}/G$ является геодезической в $\tilde{M}/G$ относительно метрики $dl = v(a)d\tilde{s}$. Этот случай представляет особый интерес и будет проанализирован ниже. Перейдем к рассмотрению эквивариантной задачи Плато.

Теорема 2 (см. [26]). Пусть G -замкнутая подгруппа в группе собственных ортогональных вращений $\mathbf{R}^n$, т. е. $G \subset \text{SO}(n)$. Пусть A^{n-2} — гладкое подмногообразие в стандартной сфере S^{n-1} , инвариантное относительно действия группы G . Тогда существует минимальная поверхность (в обобщенном смысле) X_0^{n-1} с границей A^{n-2} , инвариантная относительно действия G . Если эта G -инвариантная поверхность X_0^{n-1} единственна, то она автоматически оказывается и глобально минимальной в классе всех поверхностей с данной границей, уже не обязательно G -инвариантных.

Итак, если граница допускает нетривиальную группу симметрий G , то для нахождения абсолютно минимальной поверхности с этой границей иногда бывает достаточно найти минимальную поверхность лишь в классе G -инвариантных пленок. Различные уточнения понятия поверхности и ее границы в этой теореме см., например, в [26]. Отказ от условия, что преобразования группы G сохраняют ориентацию $\mathbf{R}^n$ (см. теорему 2), невозможен, как показывают простые примеры. Теорема 2 была распространена на случай римановых многообразий (замена $\mathbf{R}^n$ на M^n) в [31].

8. Принцип Ферма, минимальные конусы и световые пучки.

Вернемся к задаче о конусах. Предположим, что связная граница $A^{n-2} \subset S^{n-1}$ является главной орбитой (общего положения) действия компактной группы Ли G на $\mathbb{R}^n$. При проекции π на $\mathbb{R}^n/G$ граница A^{n-2} перейдет в некоторую точку $a = \pi A$. Конус CA является также G -инвариантной поверхностью и проектируется при отображении $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^n/G$ в одномерную траекторию, соединяющую a с O . Здесь O — образ особой орбиты — начала координат.

Когда конус CA глобально минимален, т. е. в каких случаях он реализует абсолютный минимум функционала объема в классе всех поверхностей с той же границей? Как мы знаем, если размерность n мала, то конус не минимален. Учитывая пункт 7, мы должны найти в $\tilde{\mathbb{R}}^n/G$ кратчайшую геодезическую, выходящую из точки a и кончающуюся на границе $\tilde{\mathbb{R}}^n/G$. Если геодезическая совпадает с траекторией, являющейся проекцией конуса, то конус будет глобально минимален. Предварительно изучим структуру $\mathbb{R}^n/G$. Рассмотрим пример: $A^{n-2} = S^{p-1} \times S^{p-1}$, где $n = 2p$. Здесь

$$A^{n-2} = \left(\sum_{i=1}^p x_i^2 + y_i^2 = 1 \right) \cap \left(\sum_{i=1}^p x_i^2 = \sum_{i=1}^p y_i^2 \right).$$

В качестве группы G возьмем границу $SO(p) \times SO(p)$, являющуюся подгруппой в $SO(2p)$. Первая компонента $SO(p)$ действует на $\mathbb{R}^p(x_1, \dots, x_p)$, а вторая — на $\mathbb{R}^p(y_1, \dots, y_p)$. Факторпространство $\mathbb{R}^{2p}/G$, очевидно, двумерно, так как группа $SO(2p)$ транзитивно действует на S^{p-1} , т. е. преобразованиями группы можно перевести любую точку сферы в любую. Так как ортогональные преобразования сохраняют длины векторов, то в качестве координат на $\mathbb{R}^{2p}/G$ можно взять пару чисел x, y — длины векторов $(x_1, \dots, x_p)$ и $(y_1, \dots, y_p)$. Поскольку $\mathbb{R}^p(x_1, \dots, x_p)/SO(p)$ изоморфно полупрямой $x \geq 0$, а $\mathbb{R}^p(y_1, \dots, y_p)/SO(p)$ — полупрямой $y \geq 0$, то $\mathbb{R}^{2p}/G$ изоморфно множеству точек двумерной плоскости, определяемому неравенствами $x \geq 0, y \geq 0$, т. е. первому квадранту. Найдем метрику $d\tilde{s}$, индуцированную на $\mathbb{R}^{2p}/G$ стандартной евклидовой метрикой ds на $\mathbb{R}^{2p}$. Рассмотрим двумерную площадку, ортогональную орбите общего положения $S^{p-1} \times S^{p-1}$ в ее произвольной точке. Ясно, что скалярное произведение, возникающее на этой площадке и позволяющее измерять расстояние между бесконечно близкими орбитами, порождается двумерной евклидовой метрикой, так как группа G действует ортогонально. Найдем функцию объема $v(a)$ на пространстве орбит. Для этого нужно вычислить $(2p-2)$ -мерный объем прямого произведения двух сфер $S^{p-1}(x) \times S^{p-1}(y)$, радиусы которых равны x и y соответственно. Так как объем $(p-1)$ -мерной сферы радиуса x равен αx^{p-1} , где α — некоторая постоянная, то $v(a) = \alpha^2 (xy)^{p-1}$. Итак, теперь мы можем вычислить метрику dl на $\mathbb{R}^{2p}/G$. Мы имеем (с точностью до постоянного множителя, отличного от нуля):

$dl^2 = (xy)^{2p-2}(dx^2 + dy^2)$. Теперь нам нужно найти локально-минимальное подмногообразие A^{2p-2} в S^{2p-1} , являющееся орбитой общего положения указанного действия G . Для этого нужно найти экстремумы функции объема $v(a)$ на $S^{2p-1}/G \subset \mathbb{R}^2/G$. Явный ее вид был вычислен ранее. Очевидно, что сфера S^{2p-1} проектируется при отображении π на дугу окружности $x^2 + y^2 = 1$, расположенную в первом квадранте. Итак, осталось найти экстремумы функции объема, ограниченной на эту окруж-

ность. Имеем: $v = (xy)^{p-1} = x^{p-1} \cdot (1 - x^2)^{\frac{p-1}{2}}$. Ясно, что функция имеет три экстремума: $x=0$, 1 , $1/\sqrt{2}$. Среди них есть только один экстремум, интересующий нас и соответствующий орбите общего положения, это $x=1/\sqrt{2}$. Итак, локально-минимальная орбита расположена на биссектрисе первого квадранта. Подведем некоторые итоги.

Фактор-пространство $\mathbb{R}^{2p}/G$ является замкнутым первым квадрантом двумерной плоскости, метрика dl^2 имеет вид $(xy)^{2p-2}(dx^2 + dy^2)$, локально-минимальное подмногообразие $A = S^{p-1} \times S^{p-1}$, вложенное в сферу S^{2p-1} , изображается точкой a , расположенной на биссектрисе первого координатного угла, конус CA изображается отрезком биссектрисы, соединяющим точку a с началом координат. Наша задача: найти геодезическую наименьшей длины, идущую из точки на границу области $\tilde{\mathbb{R}}^n/G$.

Для дальнейшего нам будет полезна следующая механическая аналогия. Функция $v(a) = n(x, y)$ является гладкой на $\tilde{\mathbb{R}}^n/G$ и равна нулю на границе. Рассмотрим двумерную сплошную прозрачную среду, заполняющую первый квадрант на евклидовой плоскости и имеющую показатель преломления $n(x, y) = c/\omega(x, y)$, где c — скорость света в пустоте, а $\omega(x, y)$ — скорость распространения света в среде в точке с координатами (x, y) . Прозрачная среда предполагается изотропной в каждой точке. Однако среда не является однородной, если коэффициент преломления не постоянен. Принцип Ферма утверждает, что световой луч, распространяющийся в среде из точки P в точку B , движется по траектории, минимизирующей при фиксированной энергии время прохождения от P до B . Далее, траектории распространения световых лучей в среде с показателем преломления $n(x, y)$ являются в то же время геодезическими конформной метрики $n(x, y)(dx^2 + dy^2)$. Верно и обратное, поэтому для описания поведения пучка световых лучей, выпущенных из какой-то точки-источника, достаточно описать структуру пучка геодезических, исходящих из этой точки во всех направлениях. Рассмотрим частный случай, когда среда заполняет верхнюю полуплоскость (над осью Ox) и коэффициент преломления является гладкой функцией, равной нулю на границе области, не зависящей от координаты x и монотонно возрастающей с ростом y , т. е. $n(x, y) = n(y)$, $n(0) = 0$,

$n(y_1) > n(y_2)$ при $y_1 > y_2$. Выпустим из точки, расположенной внутри среды, узкий пучок световых лучей по направлению к границе, т. е. к горизонтальной оси. Качественная картина распространения пучка показана на рис. 57. Вертикальный луч выходит на границу ортогонально оси, а близкие лучи отклоняются вправо и влево. Этим мы вскоре воспользуемся. Перейдем к интересующему нас случаю, когда $n(x, y) = (xy)^{p-1}$, $p \geq 1$. Из симметрии функции относительно перестановки координат вытекает, что биссектриса является геодезической и, сле-

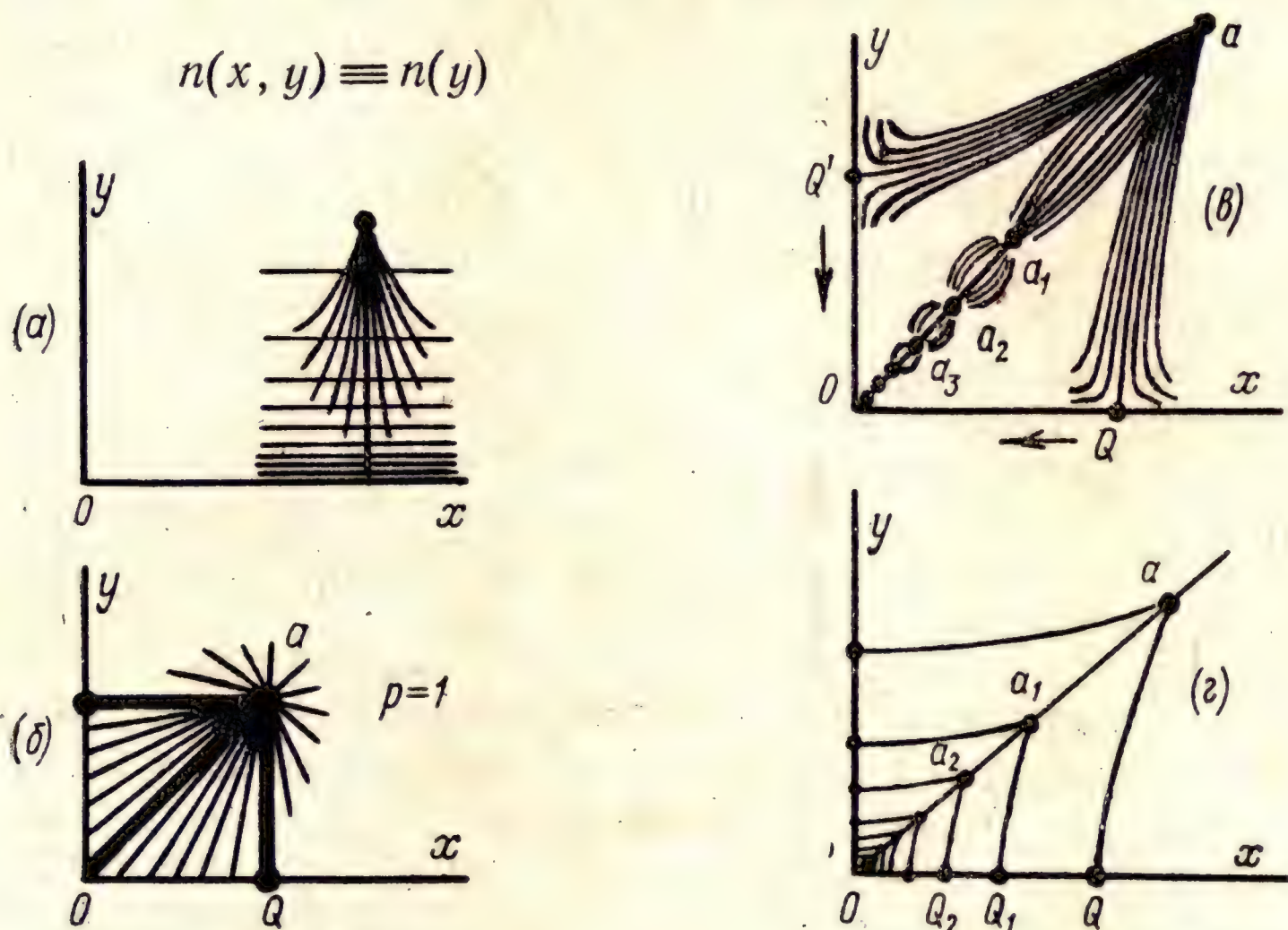


Рис. 57(а), (б)

Рис. 57(в), (г)

довательно, конус CA локально-минимален во всех своих особых точках, т. е. имеет нулевую среднюю кривизну при всех $p \geq 1$. Заметим, что это отнюдь не противоречит известной нам информации о том, что в пространствах $\mathbb{R}^n$ в малых размерностях не существует минимальных конусов, отличных от плоских дисков. Дело в том, что локальная минимальность конуса в смысле дифференциальной геометрии (т. е. минимальность относительно малых вариаций с малым носителем, сосредоточенным вблизи регулярной точки) не препятствует существованию сокращающих деформаций с достаточно большим носителем или существованию малых сокращающих деформаций в окрестности особых точек. Именно это и происходит в малых размерностях. Например, при $p=1$ ситуация тривиальна и знакома нам. Здесь $A^0 = S^0 \times S^0 \subset S^1 \subset \mathbb{R}^2$ и конус CA состоит из двух диаметров, пересекающихся в начале координат под прямым углом. Хотя каждый из них является геодезической, тем не менее в вершине одномерного конуса существует сокращающая деформация, приводящая к распаду четырехкратной

точки в сумму двух трехкратных точек. Истинно минимальная одномерная поверхность состоит здесь из вертикальных отрезков (рис. 55). С точки зрения последней модели на двумерной плоскости метрика dl при $p=1$ имеет вид $dx^2 + dy^2$, и картина распространения лучей из точки a , показанная на рис. 57(б), совпадает с пучком прямых, выходящих из точки a . Ясно, что биссектриса не является кратчайшей геодезической, а таковой является отрезок aQ , ортогонально встречающий ось Ox . Он и порождает одномерный цилиндр (рис. 55). Именно сейчас мы, наконец, получаем интерпретацию для траекторий, показанных в последнем столбце рис. 55. Верхняя половина кривых, оказывается, изображает наименьшую геодезическую, соединяющую точку-источник a с границей области $\mathbb{R}^n/G$, совпадающей с частью плоскости.

Пусть $p > 1$. Для простоты будем считать p непрерывным параметром. Несмотря на то что первоначально число p изображало целочисленную размерность, тем не менее с точки зрения анализа геодезических метрики dl нет необходимости ограничиваться лишь целыми значениями параметра. Более того, интересующий нас качественный скачок, изменение картины геодезических, происходит как раз при некотором критическом значении p_0 , которое не является целым.

Оказывается, с ростом p вертикальный отрезок aQ , изображавший наименьшую геодезическую при $p=1$, начинает прогибаться по направлению к началу координат и, что особенно важно, остается наименьшей геодезической, соединяющей точку a с границей области. Точка Q начинает монотонно двигаться по направлению к точке O . В силу симметрии метрики dl относительно перестановки координат x, y такие же события происходят и над биссектрисой. При $p > 1$ ситуация усложняется. Качественная картина распространения светового пучка, исходящего из a , показана на рис. 57(в) при $p > 1$. Интересно, что на отрезке биссектрисы от точки a до начала координат возникает (при значениях p близких к единице) бесконечное число ярко освещенных точек $a_1, a_2, a_3, \dots$, в которых собираются близкие геодезические, организованные в узкий пучок, выпущенный вдоль биссектрисы по направлению к началу координат. Первой такой точкой будет a , где соберутся все геодезические, бесконечно близкие к биссектрисе. Такие точки называются сопряженными. Прямое вычисление показывает, что таких точек бесконечно много. Качественное рассуждение, предсказывающее именно такую картину распространения светового пучка, следует из сделанного выше замечания о характере преломления световых лучей, движущихся к границе среды по направлению падения коэффициента преломления. Ясно, что это объясняет поведение близких к траектории aQ геодезических (рис. 57(в)). Аналогичными рассуждениями с углами преломления лучей предсказывается наличие сопряженных точек вдоль биссектрисы при достаточно малых значениях

$p > 1$. При этом, правда, не удастся указать точного критического значения p_0 для размерности p , после которого ситуация меняется. Аналитические вычисления приведены, например, в [27], причем не только для рассматриваемого нами случая, но и для других примеров минимальных конусов. Ясно, что если в качестве источника взять не точку a , а точку a_i , то картина распространения светового пучка из этого источника будет тождественна изображенной на рис. 57(в). В частности, из каждой точки a_i выходит геодезическая $a_i Q_i$, отличная от биссектрисы, выходящая под прямым углом на границу и, что важно, являющаяся наименьшей геодезической, соединяющей a_i с границей (рис. 57 (г)). Так как при $k=1$ длина траектории $\tilde{X}/G$ равна объему соответствующей поверхности $X \subset M$, то при малых p ($p > 1$) конус не глобально минимален. Существование сокращающей деформации видно из рис. 57. Рассмотрим произвольную точку a_i , сопряженную с точкой a вдоль геодезической — биссектрисы. Так как в этой точке пересекаются геодезические, бесконечно близкие к биссектрисе и выпущенные из точки-источника a , то в силу теории Морса [1, 2] в окрестности a_i существует деформация, уменьшающая длину геодезической. Возникает сокращающая деформация и для соответствующей поверхности, что и доказывает отсутствие минимальности конуса. Эта деформация, имея малый носитель на $\mathbb{R}^n/G$, уже не обладает этим свойством при переходе к первоначальной поверхности в $\mathbb{R}^n$, так как построенная эквивариантная деформация должна распространяться вдоль всей орбиты, соответствующей точке a_i . В то же время можно построить сокращающую деформацию, имеющую сколь угодно малый носитель и малую амплитуду, но такая деформация будет сосредоточена в окрестности вершины конуса (рис. 57(г)). В качестве «возмущенного конуса» возьмем поверхность, инвариантную при действии группы G и изображающуюся на фактор-пространстве $\mathbb{R}^n/G$ следующей траекторией. От точки a до точки a_i , где номер i выбран достаточно большим, траектория просто совпадает с отрезком биссектрисы; далее от точки a_i до границы области траектория совпадает с кратчайшей геодезической $a_i Q_i$, выходящей на горизонтальную ось. Полная траектория, составленная из двух отрезков геодезических, имеет, очевидно, строго меньшую длину (при достаточно малых $p > 1$), чем отрезок биссектрисы от a до O , изображающий конус CA . При этом возмущении вершина конуса распадается и превращается в сферу S^{p-1} малого радиуса, являющуюся орбитой группы G . Она вырастает из точки Q_i на оси Ox . Появление «исчезающего цикла» S^{p-1} уже знакомо нам по приведенным ранее примерам (рис. 55). Построенное возмущение не является непрерывной деформацией конуса, однако обе поверхности (первоначальная и возмущенная) имеют одну и ту же границу A , например, в гомологическом смысле, хотя их топологический тип различен.

В предыдущем параграфе мы столкнулись с ситуацией, когда в функциональном пространстве всех поверхностей с данной границей существует нестабильная поверхность, т. е. являющаяся критической точкой для функционала объема и имеющая положительный индекс. Это означает, что в пространстве

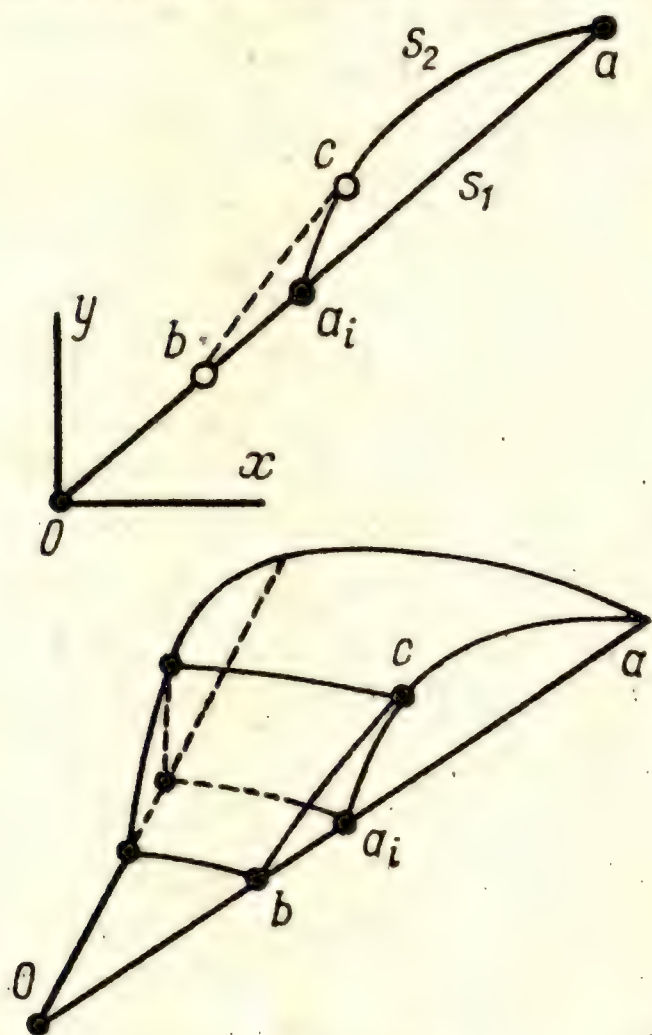


Рис. 58

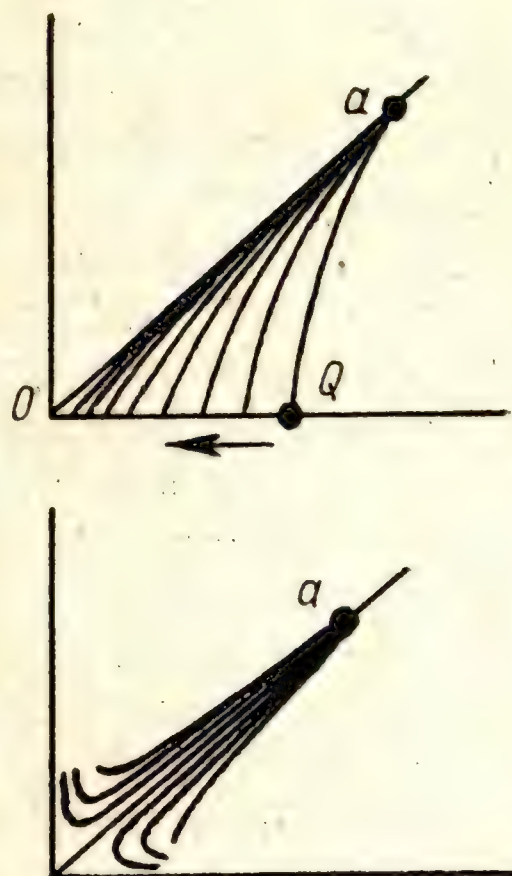


Рис. 59

всех поверхностей существуют направления, возмущения вдоль которых (сколь угодно малые по амплитуде) уменьшают объем поверхности. С геометрической точки зрения график функционала объема в малой окрестности такой экстремальной пленки имеет вид седла. Число таких независимых направлений и называется индексом. Для конусов малой размерности оказывается, что они, являясь локально-минимальными во всех своих неособых точках, являются в то же время неустойчивыми экстремальными поверхностями, индекс которых равен бесконечности. Другими словами, существует бесконечно много независимых направлений возмущения, вдоль которых объем конуса начинает уменьшаться. Такие деформации фактически уже построены нами. Пусть a_i — произвольная точка, сопряженная с точкой a вдоль биссектрисы (рис. 58). Следовательно, существует бесконечно малый поворот геодезической s_1 , соединяющей a и a_i , в близкое положение s_2 , причем при повороте точки a и a_i остаются на месте и длина поворачивающейся геодезической не меняется (с точностью до малых высших порядков). Пусть b и c — две точки, близкие к a_i и расположенные так, как показано на рис. 58. Рассмотрим бесконечно малый треугольник bca_i . Так как метрика положительно определена, то в силу неравенства треугольника длина bc меньше суммы длин ba_i и a_ic , т. е. длина ломаной $acbO$ строго мень-

ше длины ломаной $asa_i bO$. Но длина последней равна длине отрезка aO , так как длины aa_i и asa_i совпадают. Мы построили возмущение, уменьшающее длину экстремали. Соответствующее возмущение конуса, порождаемого вращением биссектрисы, см. на рис. 58. Ясно, что число таких независимых возмущений бесконечно, так как бесконечно число точек a_i , сопряженных с a .

Сделаем теперь следующий шаг и начнем увеличивать размерность p . Минимальная геодезическая aQ начинает при этом двигаться по направлению к биссектрисе. Четырехугольник $aQOQ'$ на рис. 57(в) сжимается на отрезок биссектрисы aO и, наконец, при критическом значении p_0 схлопнется на отрезок. Этот отрезок биссектрисы превратится в геодезическую наименьшей длины, соединяющую a с границей. Хотя точные вычисления затруднены (и по этой причине здесь опущены), можно утверждать, что описанные события наступают при $3,6 < p_0 < 4$. При всех $p \geq 4$ качественная картина распространения светового пучка из точки a существенно отличается от изученной нами (рис. 59). При дальнейшем росте p эта картина не меняется.

Итак, при $p \geq 4$ появляются нетривиальные глобально минимальные конусы. Первым примером оказывается конус над $S^3 \times S^3 \subset S^7 \subset \mathbb{R}^8$ при $p = 4$. Напомним, что так как при $p \geq 4$ существует единственная минимальная геодезическая, выходящая на границу области из a , и она совпадает с биссектрисой, то ее прообраз при проекции π является G -инвариантной поверхностью. В силу ее единственности из теоремы 2 следует, что она является глобально минимальной и в классе всех вариаций, уже не обязательно эквивариантных.

Прямое произведение сфер — не единственное многообразие, над которым существуют глобально минимальные конусы [25, 26, 27].

9. Проблема С. Н. Бернштейна.

Эта задача настолько тесно связана с задачей о конусах, что мы упомянем о ней именно здесь. Рассмотрим в $\mathbb{R}^n$ график гладкой функции $x_n = f(x_1, \dots, x_{n-1})$, определенной на всей гиперплоскости $\mathbb{R}^{n-1}$. График является гладким подмногообразием в $\mathbb{R}^n$. Предположим, что он локально-минимален. Вопрос: будет ли в таком случае функция f линейной, а ее график — гиперплоскостью? С. Н. Бернштейн дал положительный ответ на этот вопрос при $n = 3$. Однако проведенное им рассуждение не действовало при $n > 3$, поэтому дальше потребовалось развитие нового математического аппарата. В результате оказалось [22, 23], что окончательный ответ зависит от размерности n . Выяснилось, что при малых n любой локально-минимальный график линеен, а при больших n существуют существенно нелинейные графики, локально и даже глобально минимальные.

Первоначальным поводом к возникновению этого вопроса явились физические опыты, проведенные в XIX в. с мыльными

пленками. Давно было замечено, что если рассмотреть в $\mathbb{R}^3$ замкнутый контур, допускающий взаимно-однозначную проекцию на какую-либо фиксированную двумерную плоскость в виде выпуклой кусочно-гладкой кривой, то на контур всегда натягивается устойчивая мыльная пленка, гомеоморфная диску. Если мы начнем растягивать контур, оставляя его проекцию выпуклой, то мыльный диск начнет непрерывно деформироваться, увеличиваясь в размерах. Опыты показывают, что с увеличением контура (до пределов, обусловленных силой тяжести) мыльный диск в своей центральной части начинает уплощаться, постепенно превращаясь в поверхность, близкую к плоскости (Плато). Можно убедиться, что это уплощение начинается всегда, вне зависимости от степени сложности контура. Итак, возникает гипотеза, что при удалении контура на бесконечность мыльная пленка, затагивающая его, должна постепенно уплощаться вдали от контура и стремиться к некоторой плоскости. Таким образом, с точки зрения наблюдателя, находящегося в центре, около начала координат, мыльная пленка постепенно становится неотличимой от обычной плоскости. При этом предполагается, что наблюдатель обладает ограниченными возможностями и может видеть лишь ограниченный объем в пространстве. Редукция вопроса о минимальном графике к только что описанному выполняется на основе принципа Плато, согласно которому часть мыльной пленки, ограниченной любой замкнутой кривой, нарисованной на пленке, также оказывается минимальной поверхностью относительно этого контура. Следующий результат дает исчерпывающий ответ (Бомбьери, Де Джорджи, Джустини, Альмгрен, Саймонс).

Теорема 3 (см. [22]). Пусть $x_n = f(x_1, \dots, x_{n-1})$ — гладкая функция, определенная на всей гиперплоскости $\mathbb{R}^{n-1}$ в $\mathbb{R}^n$. Пусть ее график локально минимален, т. е. имеет нулевую среднюю кривизну. Тогда при $n \leq 8$ функция f линейна. Если же $n \geq 9$, то существуют нелинейные функции, графики которых являются не только локально, но даже глобально минимальными поверхностями.

Глобальная минимальность понимается здесь в следующем смысле. Если мы рассмотрим произвольное непрерывное возмущение с компактным носителем, т. е. оставляющее поверхность неподвижной вне компактной области, то ее объем при этом не уменьшается. С каждым минимальным графиком естественно связан некоторый минимальный конус. Оказывается, из локальной минимальности графика вытекает, что он глобально минимален. Пусть, для простоты, график проходит через начало координат. Рассмотрим шар радиуса R с центром в точке O и пусть Γ_R — часть графика, попавшая в шар (рис. 60). Изменив масштаб, т. е. выполнив преобразование подобия с коэффициентом $1/R$, сожмем шар D_R на шар D_1 радиуса 1. Поверхность Γ_R сожмется и превратится в поверхность, которую мы обозначим M_R . Она зависит от R и с ростом R как-то изменяется вну-

три шара D_1 , причем ее граница все время остается на границе шара. Рассмотрим «предел» минимальных поверхностей M_R при $R \rightarrow \infty$. Этот предел существует (в разумном смысле) и является конусом CA над некоторым $(n-2)$ -мерным подмногообразием A в сфере радиуса 1. Доказывается, что A является локально минимальным подмногообразием, причем конус глобально-минимален при фиксированной границе. В силу уже известных нам размерностных ограничений при $n \leq 7$ конус должен быть плоским диском. Отсюда выводится, что и исходный график также должен быть гиперплоскостью, что и доказывает теорему при $n \leq 7$. При $n=8$ требуется некоторое дополнительное рассуждение [22]. Построение конуса по графику можно наглядно прокомментировать, если вообразить наблюдателя, который, в отличие от предыдущего, расположен не в начале координат, а устремляется в бесконечность, и рассматривать график из бесконечно удаленной точки. Тогда все мелкие колебания графика, не являющиеся коническими, «исчезают» в начале координат.

Существование нелинейных графиков при $n \geq 9$ доказывается аналитически, хотя и с привлечением все тех же минимальных конусов [22].

10. Комплексные подмногообразия как минимальные поверхности.

Имеется широкий класс минимальных поверхностей, часто встречающихся в геометрии. Это — комплексные подмногообразия в кэлеровом многообразии M . Напомним, что комплексное многообразие называется кэлеровым, если его метрика $ds^2 = \sum_{i,j} g_{ij} dz_i d\bar{z}_j$ определяет замкнутую внешнюю дифферен-

циальную форму $\omega = \frac{i}{2} \sum_{i,j} g_{ij} dz_i \wedge d\bar{z}_j$. Здесь $z_1, \dots, z_n$ — локальные координаты, а «черта» — комплексное сопряжение.

Кэлеровыми многообразиями являются комплексное линейное пространство $\mathbb{C}^n$, проективное пространство $\mathbb{C}P^n$.

Мы уже знакомы с несколькими понятиями вариации поверхности. В простейших случаях это — непрерывная деформация, гомотопия поверхности. При алгебраическом подходе это — переход от одного цикла к другому, ему гомологичному. Сейчас мы укажем еще одно естественное понятие вариации, часто возникающее в приложениях. Пусть $X^k \subset M^n$ — гладкое компактное ориентируемое замкнутое подмногообразие. Будем говорить, что задана его бордизм-деформация, если задано $(k+1)$ -мерное гладкое компактное ориентируемое подмногообразие Z с краем $\partial Z = X \cup (-Y)$, где Y — некоторое гладкое компактное ориентируемое k -мерное подмногообразие в M . Здесь через $-Y$ обозначено многообразие Y с обратной ориентацией. Многообразие Y назовем бордизм-вариацией исходного подмногообразия X . В случае некомпактного подмногообразия X

будем говорить, что задана его бордизм-деформация, если в M задано подмногообразие Y , совпадающее с X вне некоторой компактной области, и, кроме того, задано $(k+1)$ -мерное подмногообразие Z с кусочно-гладким краем $\partial Z \subset X \cup (-Y)$ (рис. 61). Ясно, что с точки зрения гомологий два цикла $[X]$

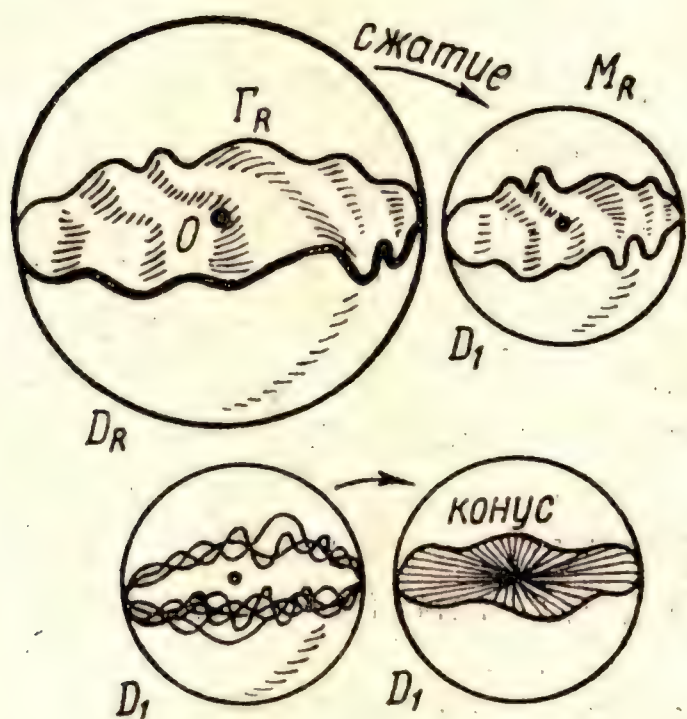


Рис. 60

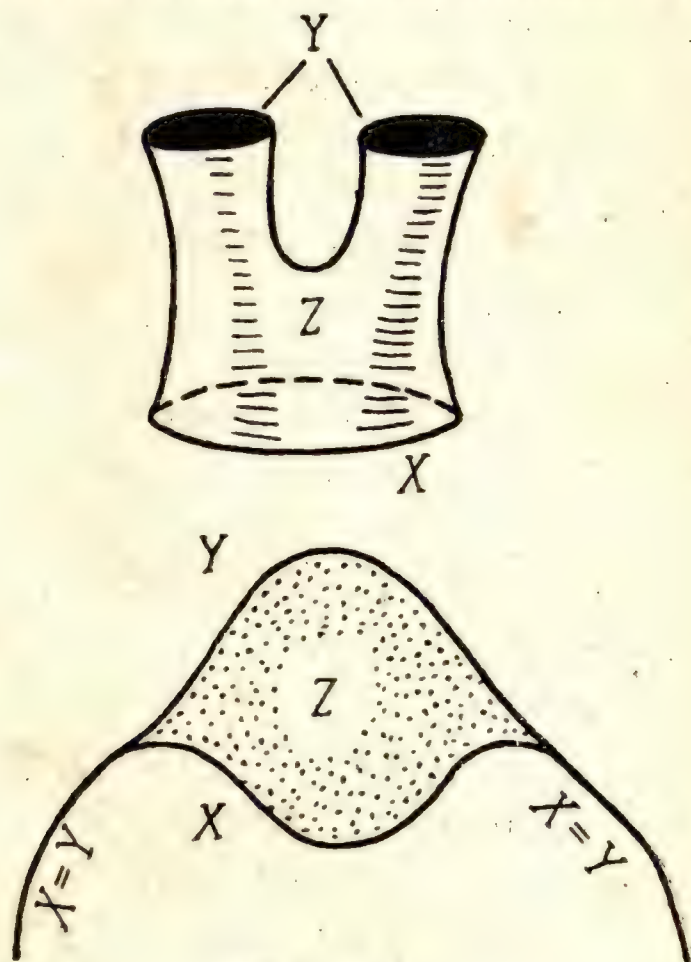


Рис. 61

и $[Y]$, реализуемые подмногообразиями X и Y , гомологичны. Однако введенное выше отношение эквивалентности, называемое иногда бордантностью, более сильно, чем отношение гомологичности: гомологичные подмногообразия могут оказаться не бордантными.

Теорема 4 (см. [9]). Пусть M — кэлерово многообразие и X — комплексное подмногообразие комплексной размерности k . Рассмотрим всевозможные вещественные бордизм-деформации этого подмногообразия в объемлющем многообразии. Пусть Y — произвольная бордизм-вариация подмногообразия X . Тогда объем подмногообразия X не превышает объема подмногообразия Y , если оба они компактны, т. е. выполнено неравенство $\text{vol}_{2k} X \leq \text{vol}_{2k} Y$. Если же подмногообразия X и Y некомпактны, то аналогичное неравенство имеет место для объемов тех областей на X и на Y , где X и Y различны. Более того, если объемы X и Y совпадают, то подмногообразие также является комплексным.

Например, в CP^n глобально минимальными являются все комплексные проективные подпространства CP^k , стандартно вложенные в CP^n . Глобально минимальными являются и алгебраические комплексные поверхности в C^n или в CP^n , хотя поверхности могут иметь особенности.

11. Интегральные потоки.

Мы познакомились с некоторыми понятиями поверхности и ее границы, начиная от представлений XVIII—XIX вв. и кон-

чая алгебраическими определениями гомологической границы. На этом список различных подходов к таким фундаментальным понятиям, как поверхность и ее граница, не исчерпывается. Чрезвычайно плодотворным оказался язык, разработанный Федерером, Флемингом, Альмгреном и др. [9] и позволивший доказать многие из результатов, вкратце затронутых в предыдущих пунктах. Это — язык потоков. Мы лишь кратко коснемся существа этого понятия, отсылая за подробностями к [9—11, 22, 23, 25, 26, 30].

Вероятно, одним из предшественников понятия интегрального потока явился процесс интегрирования δ -функции Дирака. Если $f(x)$ — функция, заданная на некоторой области определения, то мы можем рассмотреть δ -функцию, сосредоточенную в некоторой точке x_0 из этой области, и представить значение функции $f(x_0)$ в этой точке в виде результата интегрирования функции $f(x)\delta(x - x_0)$ по области определения: $f(x_0) = \int f(x)\delta(x - x_0)dx$. При желании можно смотреть на это как на процесс интегрирования нульмерной формы $f(x)$ по нульмерному подмногообразию x_0 в области определения. Меняя точку x_0 , мы меняем и результат интегрирования, если функция f не постоянная. Попытка расширить это определение и заменить нульмерное подмногообразие на подмногообразие произвольной размерности k и приводит к понятию интегрального потока. Один из вариантов такого расширения был предложен де Рамом. Здесь же мы приведем определение Флеминга и Федерера. Начнем с понятия сглаживаемых подмножеств. Будем говорить, что некоторое k -мерное ограниченное подмножество X в римановом многообразии M является k -сглаживаемым (или k -спрямляемым), если оно измеримо в смысле k -мерной меры Хаусдорфа [7, 20], и, кроме того, подмножество X может быть сколь угодно близко аппроксимировано компактным ориентируемым гладким k -мерным подмногообразием в M . Понятие близкой аппроксимации уточняется так: для любого сколь угодно малого положительного числа ε должно найтись такое гладкое k -мерное компактное ориентируемое подмногообразие $Y \subset M$, что $\text{vol}_k(X \setminus Y) \cup (Y \setminus X) < \varepsilon$, т. е. оно отличается от X на множестве точек малой меры. Здесь vol_k обозначает меру Хаусдорфа, совпадающую с обычным k -мерным римановым объемом для «хороших» множеств (подмногообразий). Можно требовать выполнения другого неравенства: $\text{vol}_k(X \cup Y) \setminus (X \cap Y) < \varepsilon$, что не влияет на дальнейшие построения. Ясно, что запас k -сглаживаемых подмножеств достаточно велик. Он включает, в частности, все конкретные мыльные пленки, возникающие в физических опытах. Одной из характеристик класса k -сглаживаемых множеств является то, что они — наиболее удобные подмножества, по которым можно естественно интегрировать k -мерные внешние гладкие дифференциальные формы. Фиксируем k , и пусть нам заданы: внешняя k -мерная форма ω и k -мерное ориентируемое компактное

подмногообразии X . Тогда однозначно определен интеграл формы по подмногообразию: $\int_X \omega$ [1, 5]. Эта операция распространяется и на класс многообразий X , погруженных в объемлющее многообразие M , т. е. допускаются самопересечения. При этом некоторые листы многообразия M могут слипаться с точки зрения M , однако при интегрировании ω по X мы, естественно, учитываем все точки X . Если образ X содержит участ-

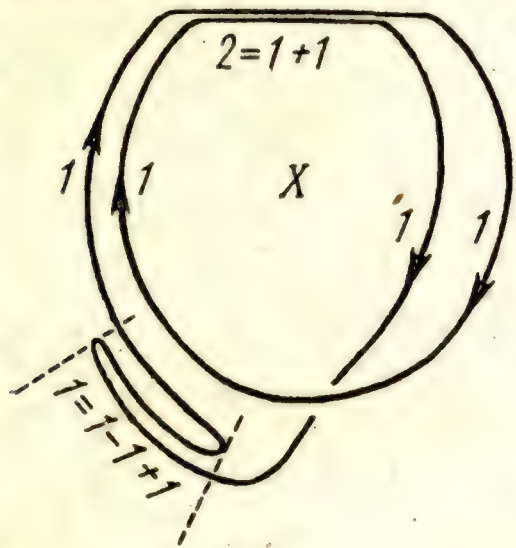


Рис. 62

ки склейки, слипания нескольких листов, то такой участок учтется при интегрировании с некоторой кратностью (рис. 62). Цифрами указана кратность листов. Можно показать, что если X является k -сглаживаемым, то для почти всех его точек x существует единственная аппроксимирующая касательная плоскость T_x в точке x к X . Поскольку k -сглаживаемое множество сколь угодно близко аппроксимируется k -мерным подмногообразием, то можно распространить операцию интегрирования с

класса погруженных подмногообразий на класс k -сглаживаемых подмножеств. Мы опустим технические детали и сразу будем считать, что определено интегрирование любой k -мерной формы по любому ориентируемому k -сглаживаемому подмножеству X в M .

Мы получаем непрерывный линейный функционал $\|X\|$, определенный на пространстве $\Lambda^k M$ всех внешних k -мерных гладких дифференциальных форм на M и отображающий его в $\mathbb{R}^1$. Нужно проинтегрировать каждую форму ω по k -сглаживаемому подмножеству X , т. е. $\|X\| : \omega \rightarrow \int_X \omega$. В этом смысле функ-

ционал $\|X\|$ и выступает в роли δ -функции, но только вместо одной точки x_0 мы рассматриваем k -мерное подмножество X . Оказывается, $\|X\|$ — неплохая «модель» множества X .

Дадим несколько важных определений. Поток размерности k в многообразии M называется непрерывный линейный функционал, отображающий пространство k -мерных внешних форм $\Lambda^k M$ в $\mathbb{R}$. Сглаживаемым потоком размерности k в многообразии M называется функционал $\|X\|$, где X есть k -сглаживаемое подмножество в M . Дадим определение границы потока. Пусть $T : \Lambda^k M \rightarrow \mathbb{R}$ — некоторый поток и $d : \Lambda^{k-1} M \rightarrow \Lambda^k M$ — обычный оператор внешнего дифференцирования. Границей ∂T потока T называется поток $\partial T : \Lambda^{k-1} M \rightarrow \mathbb{R}$, задаваемый так: $\partial T(\alpha) = T(d\alpha)$, где $\alpha \in \Lambda^{k-1} M$. Это определение обобщает понятие обычной границы многообразия. Связь между этими определениями осуществляется через классическую формулу Стокса. В самом деле, пусть поток T является сглаживаемым, т. е. $T = \|X\|$, и пусть X есть k -мерное ориентируемое подмногообразие с краем ∂X . Тогда $\partial \|X\|(\alpha) = \|X\|(d\alpha)$, т. е. $\int_{\partial X} \alpha = \int_X d\alpha$.

Последнее равенство и является формулой Стокса [1, 5]. Итак, если поток является k -сглаживаемым и порожден многообразием с краем, то его граница интерпретируется как функционал, порожденный обычной границей подмногообразия X . Конечно, в общем случае такая непосредственная связь между границей потока и «границей множества X » отсутствует. Вернемся снова к понятию потока.

Определение 1. Поток T называется интегральным, если он представим в виде суммы с положительными или отрицательными целыми коэффициентами k -сглаживаемых потоков (в случае, когда сумма бесконечна, предполагается, что этот ряд сходится) и если его граница ∂T обладает тем же свойством.

Итак, интегральный поток T имеет вид: $T = \sum_i a_i \|X_i\|$, где

a_i — целые числа, а X_i — k -сглаживаемые подмножества в M . Носителем интегрального потока T называется объединение множеств X_i , т. е. $\text{supp } T = \bigcup_i X_i$. Носитель потока может

быть определен также как наименьшее множество, вне которого поток аннулируется (при интегрировании форм по множествам, не задевающим носитель). Роль интегральных потоков заключается в первую очередь в том, что в их терминах удастся доказывать теоремы существования минимальных потоков, причем носители потоков оказываются в этом случае почти всюду гладкими и даже аналитическими подмногообразиями [9]. При этом обычно минимизируется так называемая масса потока, вычисляющаяся для потока $T = \sum_i a_i \|X_i\|$ так:

$$m(T) = \sum_i |a_i| \text{vol}_k X_i.$$

Многие из перечисленных ранее теорем (например, существование G -инвариантной минимальной поверхности) формулируются и доказываются в терминах интегральных потоков. Интегральные потоки размерности k в многообразии M образуют абелеву группу по сложению, которую обозначим $I_k(M)$. Пусть, для простоты, $M = \mathbb{R}^n$, тогда будем писать просто I_k . Рассмотрим в $\mathbb{R}^n$ два подмногообразия A и B , причем $A \subset B$. Построим цепной комплекс:

$$\Omega(B, A) = \bigoplus_{k=0}^n I_k[B, A],$$

где $I_k[B, A] = I_k \cap \{T : \text{supp } T \subset B, \text{supp } \partial T \subset A\}$. Граничный оператор ∂ определен так: $\partial : I_k[B, A] \rightarrow I_{k-1}[B, A]$. Поступая по схеме § 1 гл. 1, мы можем вычислить гомологии этого комплекса, которые обозначим $H_i(\Omega(B, A), \partial)$. Имеет место важная теорема [9], связывающая гомологии комплекса интегральных потоков с обычными гомологиями пары (B, A) с це-

лыми коэффициентами, а именно имеет место изоморфизм $H_i(\Omega(B, A), \partial) \cong H_i(B, A, \mathbb{Z})$. В частности, каждый i -мерный класс гомологий $\sigma \in H_i(B, A, \mathbb{Z})$ представляется (с помощью изоморфизма α) i -мерным интегральным потоком наименьшей массы, т. е. его масса наименьшая по сравнению с массами других потоков из данного гомологического класса [9, 10].

12. Бордизмы и классическая многомерная задача Плато.

Рассмотрим риманово многообразие M , выделим в нем гладкое компактное замкнутое (т. е. не имеющее края) $(k-1)$ -мерное подмногообразие A — «контур». Пусть существует хотя бы одно k -мерное гладкое многообразие W с краем

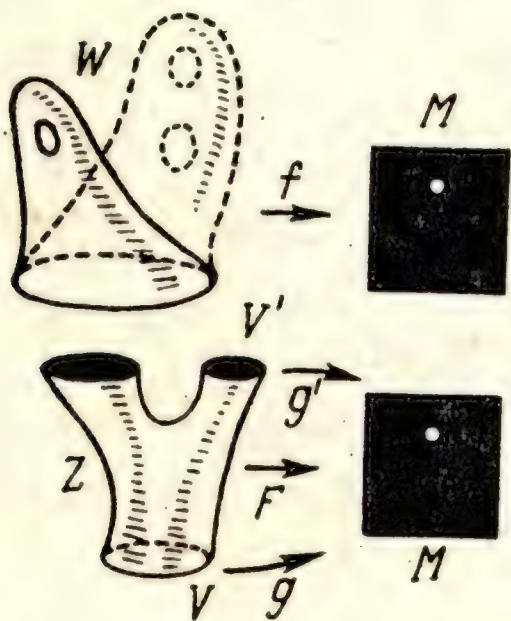


Рис. 63

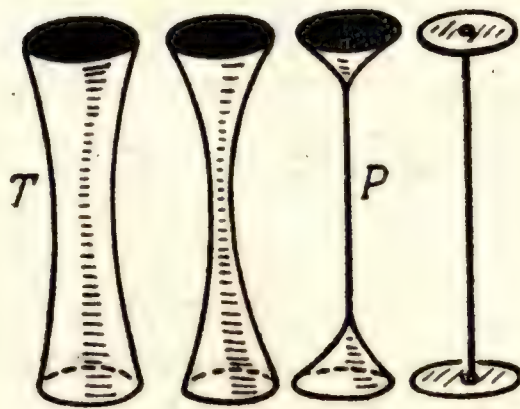


Рис. 64

А. Рассмотрим всевозможные пары вида (W, f) , где W — компактное многообразие с краем ∂W , гомеоморфным A , а $f: W \rightarrow M$ — непрерывное (или кусочно-гладкое) отображение, тождественное на крае ∂W .

Задача А. Можно ли среди всех пар вида (W, f) , описанных выше, найти такую пару (W_0, f_0) , чтобы отображение f_0 или пленка $X_0 = f_0(W_0)$, т. е. образ W_0 в M , обладали бы разумными свойствами минимальности, в частности, чтобы $\text{vol}_k X_0 \leq \text{vol}_k X$, где $X = f(W)$ — любая другая пленка из указанного класса?

Задача А является задачей нахождения абсолютного минимума функционала объема по всем пленкам, заклеивающим данный контур A (рис. 63). Следующая задача уточняет задачу о нахождении абсолютного минимума в классе всех бордизм-вариаций (см. п. 10 настоящего параграфа) данного подмногообразия $V \subset M$.

Задача Б. Пусть (V, g) — пара, где V — компактное замкнутое k -мерное многообразие, g — его непрерывное (или кусочно-гладкое) отображение в M , $X = g(V)$ — образ V в M . Мы скажем, что пара (V', g') является бордизм-вариацией пары (V, g) , если существует компактное многообразие Z с краем $\partial Z = V \cup (-V')$ и непрерывное отображение $F: Z \rightarrow M$ такое, что его ограничения на компоненты края V и V' совпа-

дают с заданными ранее отображениями g и g' соответственно. Можно ли среди всех бордизм-вариаций найти такую пару (V_0, g_0) , чтобы образ $X_0 = g_0(V_0)$ обладал разумными свойствами минимальности, в частности, чтобы $\text{vol}_k X_0 \leq \text{vol}_k X$, где X — любая другая пленка из указанного класса (рис. 63)?

В обеих задачах топологический тип пленки может меняться, хотя и не совсем произвольно. При решении двумерной задачи Плато игнорируется следующий эффект, начинающий играть существенную роль в больших размерностях. Рассмотрим знакомый нам эффект перестройки катеноида в объединение двух дисков при растяжении граничного контура. На рис. 64 показан контур, состоящий из двух окружностей и затянутый пленкой $X_t = f_t(W)$, стремящейся занять положение с наименьшей площадью (поверхностной энергией). Ясно, что в некоторый момент происходит схлопывание, перетяжка пленки и вместо трубки T появится отрезок P , соединяющий верхний и нижний диски. В двумерном случае от него легко избавиться, удалив его и не потеряв при этом заклеивающих свойств пленки: ее остаток (два диска) по-прежнему заклеивает границу. В случае $k > 2$, когда в качестве границы выступает $(k - 1)$ -мерное подмногообразие, возникновение ситуации, аналогичной описанной, резко усложняет задачу минимизации и поиска минимального решения. По мере того как k -мерный объем деформирующейся пленки $X_t = f_t(W)$, $0 \leq t \leq 1$, стремится к минимуму, пленка стремится занять соответствующее минимальное положение, и в ней могут возникнуть склейки, аналогичные описанным. Другими словами, отображение $f_1: W \rightarrow M$, гомотопное исходному, может понижать размерность на некоторых открытых в W подмножествах. Это приводит к появлению в образе $X_1 = f_1(W)$ кусков P размерностей меньших, чем k . В отличие от двумерного случая все такие куски нельзя, вообще говоря, ни отбросить, ни непрерывно отобразить в «массивную» k -мерную часть $X^{(k)}$ пленки X_1 . Напомним, что перед нами стоит задача нахождения минимальной пленки в классе пленок, допускающих непрерывную параметризацию с помощью многообразий W , поэтому при любом варианте устранения маломерных кусков P мы должны гарантировать, что пленка, получившаяся после такой перестройки, будет по-прежнему допускать непрерывную параметризацию с помощью многообразия. Подробное изучение всех возможных вариантов перестроек показывает, что (хотим мы этого или нет) мы не можем игнорировать куски малых размерностей, автоматически возникающие при минимизации функционала объема, и вынуждены считаться с ними. Подробности см. в [27]. Можно было бы в целях упрощения задачи ограничиться пока рассмотрением лишь k -мерного функционала объема, с точки зрения которого все маломерные «добавки» несущественны, поскольку имеют нулевой k -мерный объем. Однако оказывается [27], что даже в этом случае решение задачи Плато требует получения обшир-

ной информации о поведении маломерных кусков пленки. Так или иначе мы снова приходим к задаче одновременного изучения как функционала vol_k , так и всех остальных функционалов vol_s (при s меньших, чем k), определенных на маломерных частях пленки. Следовательно, математическая природа многомерной задачи Плато вынуждает нас ввести: а) стратифицированные поверхности $X = X^{(k)} \cup X^{(k-1)} \cup \dots$, где каждое подмножество (страт) $X^{(s)}$ является s -мерной «поверхностью» в M , имеющей в каждой своей точке размерность s ; б) стратифицированный объем $SV(X) = (\text{vol}_k X^{(k)}, \text{vol}_{k-1} X^{(k-1)}, \dots)$, изображаемый вектором, каждая компонента которого равна объему соответствующего страта (в каждой размерности).

Нахождение минимальной поверхности X означает доказательство существования стратифицированной поверхности, стратифицированный объем SV которой — наименьший. Под минимизацией вектора стратифицированного объема понимается следующее: сначала нужно минимизировать его первую координату (т. е. k -мерный объем), затем, фиксируя это минимальное значение первой координаты, минимизировать вторую координату, затем, фиксируя минимальные значения первых двух координат, минимизировать третью и т. д. Для реализации этой программы нужен новый язык, на котором можно точно поставить задачу. Минимизация вектора объемов SV понимается лексикографически.

Определение 2. Замкнутое ориентированное многообразие V^{k-1} называется бордантным нулю, $V \sim 0$, если существует компактное ориентированное многообразие W^k , край которого ∂W диффеоморфен с сохранением ориентации многообразию V . Два замкнутых ориентированных многообразия V_1 и V_2 называются бордантными, $V_1 \sim V_2$, если их несвязное объединение $V_1 \cup (-V_2)$ бордантно нулю.

Отношение бордантности является отношением эквивалентности. Множество Ω_{k-1} классов эквивалентности является абелевой группой, в которой сложение индуцировано несвязным объединением многообразий. Прямую сумму групп $\bigoplus_k \Omega_{k-1}$ обычно обозначают через Ω_* . В этой сумме определена структура кольца, индуцированная прямым произведением многообразий. Большую роль в теории вариационных задач играют также неориентированные бордизмы $N_* = \bigoplus_k N_{k-1}$. Для построения групп N_{k-1} используются все замкнутые многообразия, без ограничения на ориентируемость.

Определение 3. Ориентированным сингулярным многообразием пространства M назовем пару (V, f) , где V^{k-1} — компактное ориентированное многообразие, а f — непрерывное отображение $V \rightarrow M$.

Как и выше, определяются бордантные сингулярные многообразия. Множество классов эквивалентности снабжается структурой абелевой группы и обозначается $\Omega_{k-1}(M)$. Эта группа называется $(k-1)$ -мерной группой ориентированных

бордизмов пространства M . Аналогичное построение, но без учета ориентируемости, приводит нас к определению групп $N_{k-1}(M)$ неориентированных бордизмов. Вернемся к задаче Плато. Задачи А и Б (см. выше) допускают следующие формулировки (А. Т. Фоменко):

Задача А. Можно ли среди всех компактов X , содержащих A и обладающих тем свойством, что сингулярный бордизм (A, i) эквивалентен нулю в X (где $i: A \rightarrow X$ — вложение), найти компакт X_0 , обладающий свойствами минимальности?

Так как тождественное отображение $e: A \rightarrow A$ определяет некоторый элемент $\sigma \in \Omega_{k-1}(A)$, то введенный класс пленок-компактов X характеризуется тем, что $i_*\sigma = 0$, где $i_*: \Omega_{k-1}(A) \rightarrow \Omega_{k-1}(X)$ — гомоморфизм, индуцированный вложением $i: A \rightarrow X$.

Задача Б. Можно ли среди всех сингулярных многообразий (V, g) , $g: V \rightarrow M$, бордантных (эквивалентных) данному сингулярному многообразию (V', g') , $g': V' \rightarrow M$, найти такое сингулярное многообразие (V_0, g_0) , чтобы пленка $X_0 = g_0(V_0)$ обладала свойствами минимальности? Иначе говоря, можно ли среди всех представителей (V, g) данного класса бордизмов $\alpha \in \Omega_k(M)$ найти такой (V_0, g_0) , что пленка $X_0 = g_0(V_0)$ минимальна в M ?

Как и в п. 2, для аккуратного построения теории минимизации нужно уметь вычислять группы бордизмов не только для многообразий и клеточных комплексов, но и для поверхностей с особенностями, поскольку минимальные пленки содержат, вообще говоря, достаточно сложные особые точки, заполняющие подмножества нулевой меры. Это расширение понятия группы бордизмов на более широкий класс компактных подмножеств осуществляется при помощи так называемого спектрального процесса, как и в случае построения спектральных гомологий [13—16, 27]. Теперь мы подготовлены для формулировки теоремы, решающей задачу Плато в терминах образов многообразий с краем, т. е. в постановке, максимально близкой к классической двумерной задаче. Фиксируем в $\mathbf{R}^n$ «контур» A^{k-1} , в качестве которого возьмем для простоты замкнутое гладкое подмногообразие. Пусть $A(\sigma)$ — класс всех измеримых компактных подмножеств X в $\mathbf{R}^n$, которые содержат A и которые аннулируют бордизм $\sigma = (A, e)$, где $e: A \rightarrow A$ — тождественное отображение. Другими словами, сингулярное многообразие (A, i) , где $i: A \rightarrow X$ — вложение, бордантно нулю в пленке X , т. е. $i_*\sigma = 0$, где i_* — гомоморфизм, отображающий $(k-1)$ -мерные бордизмы A в $(k-1)$ -мерные бордизмы пленки X . Через vol_k обозначим либо риманов объем пленки X , либо ее k -мерную хаусдорфову меру, когда пленка содержит достаточно сложные особенности. Автором книги доказана следующая

Теорема 5 (см. [13—15]). (1) Пусть $\{X\}_k$ — класс всех компактов X , $A \subset X \subset \mathbf{R}^n$ таких, что $X \in A(\sigma)$ и $\text{vol}_k X = d_k = \inf_Y \text{vol}_k Y, Y \in A(\sigma)$. Тогда класс $\{X\}_k$ непуст, $d_k > 0$ и каждый

компакт $X \in \{X\}_k$ содержит однозначно-определенное k -мерное подмножество T^k (т. е. имеющее в каждой своей точке размерность k), $T^k \subset X \setminus A$, такое, что $A \cup T^k$ является компактом в $\mathbb{R}^n$, множество T^k содержит некоторое «особое» подмножество Z_k (которое может быть пусто), где $\text{vol}_k Z_k = 0$ и $T^k \setminus Z_k$ является гладким минимальным k -мерным подмногообразием в $\mathbb{R}^n$ (в действительности $T^k \setminus Z_k$ даже аналитическое подмногообразие), причем многообразие $T^k \setminus Z_k$ открыто и всюду плотно в множестве T^k и $\text{vol}_k T^k = \text{vol}_k X = d_k > 0$.

(2) Далее, если $\{X\}_{k-1}$ — класс всех компактов X , $A \subset X \subset \mathbb{R}^n$, таких, что $X \in A(\sigma)$, $X \in \{X\}_k$ и $\text{vol}_{k-1} X \setminus T^k = d_{k-1} = \inf_Y \text{vol}_{k-1} Y \setminus T^k$, $Y \in \{X\}_k$, то класс $\{X\}_{k-1}$ непуст, а в том случае, когда $d_{k-1} > 0$, каждый компакт $X \in \{X\}_{k-1}$ содержит однозначно-определенное $(k-1)$ -мерное подмножество T^{k-1} (имеющее в каждой своей точке размерность $k-1$), $T^{k-1} \subset X \setminus (A \cup T^k)$, причем подмножество $A \cup T^k \cup T^{k-1}$ является компактом в $\mathbb{R}^n$, T^{k-1} содержит некоторое «особое» подмножество Z_{k-1} (которое может быть пусто), где $\text{vol}_{k-1} Z_{k-1} = 0$ и $T^{k-1} \setminus Z_{k-1}$ является гладким минимальным $(k-1)$ -мерным подмногообразием в $\mathbb{R}^n$, всюду плотным в T^{k-1} , причем $\text{vol}_{k-1} T^{k-1} = \text{vol}_{k-1} X \setminus (A \cup T^k) = d_{k-1} > 0$. Если же $d_{k-1} = 0$, то положим $T^{k-1} = \emptyset$.

(3), (4), ... И так далее, вниз по размерностям.

Следствие 1. В классе поверхностей $A(\sigma)$ всегда существует глобально минимальная поверхность $X_0 = A \cup T^k \cup T^{k-1} \cup \dots$, стратифицированный объем которой $SV(X_0) = (\text{vol}_k T^k, \text{vol}_{k-1} T^{k-1}, \dots)$ является наименьшим (в указанном смысле) среди стратифицированных объемов других поверхностей $X \in A(\sigma)$. Минимальная пленка X_0 является глобально минимальной во всех размерностях. Каждый из стратов X_0 является (за исключением, быть может, множества точек меры нуль) аналитическим минимальным подмногообразием в $\mathbb{R}^n$.

Аналогичная теорема имеет место и в задаче Б, т. е. в задаче реализации бордизма с помощью пленки, минимальной во всех размерностях. Эти теоремы являются частными случаями общей теоремы существования глобально минимальных во всех размерностях стратифицированных поверхностей, доказанной автором в [13—15] для широкого класса многомерных вариационных задач. Итак, в классе поверхностей с фиксированным краем A и допускающих непрерывную параметризацию в виде образов гладких многообразий с краем A , всегда существует глобально минимальная поверхность с указанными свойствами.

13. Существование минимума в каждом гомотопическом классе.

Рассмотрим фиксированную границу $A \subset M$, где A есть $(k-1)$ -мерное замкнутое многообразие. Рассмотрим пленки, имеющие A своей границей и получающиеся друг из друга непрерывной деформацией.

Задача A' . Можно ли среди всех пар вида (W, f) , где

W — некоторое фиксированное многообразие с краем A , а $f: W \rightarrow M$ — всевозможные непрерывные (или кусочно-гладкие) отображения, гомотопные некоторому фиксированному отображению f' и тождественные на крае A (т. е. совпадающие с фиксированным гомеоморфизмом A на себя), найти такую пару (W, f_0) , чтобы отображение f_0 или пленка $X_0 = f_0(W)$ обладали бы свойствами минимальности, в частности, чтобы $\text{vol}_k X_0 \leq \text{vol}_k X$, где $X = f(W)$ — любая пленка из данного гомотопического класса?

Задача Б'. Можно ли среди всех отображений $g: V \rightarrow M$ (где V — фиксированное многообразие), гомотопных исходному отображению $f: V \rightarrow M$, найти такое отображение g_0 , которое обладало бы свойством минимальности, в частности, чтобы $\text{vol}_k g_0(V) \leq \text{vol}_k g(V)$, где g — любой представитель указанного класса отображений?

Из решения задач А и Б (см. выше) не следует решение задач А' и Б' (как и обратно). С точки зрения объемлющего многообразия M задачи А и Б — это задачи на отыскание абсолютного минимума, а задачи А' и Б' — это задачи на отыскание относительных минимумов в классе всех пленок, т. е. абсолютных минимумов в каждом гомотопическом классе по отделимости. Дао Чонг Тхи разработал теорию стратифицированных потоков, являющуюся функциональным аналогом теории стратифицированных поверхностей, описанных выше [13—15]. Эти стратифицированные потоки были названы в [28] мультиварифолдами. Задача А' была сформулирована так: можно ли среди всех локально липшицевых отображений $g: W^k \rightarrow M$, гомотопных некоторому исходному отображению $f: W^k \rightarrow M$ и таких, что $g|_{\partial W} = f|_{\partial W}$, найти отображение g_0 , минимизирующее старший функционал k -мерного объема (при этом ∂W может быть пусто). При таком подходе минимизируется лишь старший объем vol_k , а поведение объемов меньших размерностей не контролируется. В [28] установлено существование локально липшицева отображения $g_0: W \rightarrow M$ (в терминах потоков), минимизирующего функционал старшего k -мерного объема.

14. Случай отсутствия решения задачи Дирихле для уравнения минимальной поверхности большой коразмерности.

Мы познакомились со свойствами вещественнозначных функций, графики которых являются минимальными поверхностями коразмерности один в $\mathbf{R}^{n+1}$. Существенно меньше известно о свойствах вектор-функций, графики которых являются минимальными поверхностями большей коразмерности. Оказывается, с увеличением коразмерности свойства графиков усложняются и появляются новые эффекты, отсутствующие в случае коразмерности один. Некоторые примеры мы приведем в настоящем пункте. При этом мы будем в основном опираться на работу [30] Лоусона и Оссермана.

Пусть B — открытое подмножество в $\mathbf{R}^n(x_1, \dots, x_n)$ и пусть $F: B \rightarrow \mathbf{R}^{n+k}$ — гладкое погружение. Отображение F определяет

минимальную поверхность $F(B)$ в $\mathbf{R}^{n+k}$ в том и только в том случае, когда F удовлетворяет следующей системе дифференциальных уравнений:
$$\sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(\sqrt{g} g^{ij} \frac{\partial F}{\partial x_j} \right) = 0.$$
 Здесь g_{ij} —

компоненты метрики, индуцированной на $F(B)$ объемлющей евклидовой метрикой, т. е. $g_{ij} = \langle \partial F / \partial x_i, \partial F / \partial x_j \rangle$. Далее g^{ij} — компоненты матрицы, обратной к матрице $\|g_{ij}\|$, наконец, g — определитель $\|g_{ij}\|$. Вывод этого векторного уравнения из уравнения Эйлера для функционала объема мы предоставляем читателю. Будем говорить, что погружение F области B в $\mathbf{R}^{n+k}$ является непараметрическим, если F имеет вид $F(x) = (x, f(x))$, где f — отображение (вектор-функция) $f: B \rightarrow \mathbf{R}^k$. Тогда указанная выше система уравнений приобретает вид:

$$(*) \sum_{i=1}^n \frac{\partial}{\partial x_i} (\sqrt{g} g^{ij}) = 0, \quad 1 \leq j \leq n; \quad \sum_{i,j=1}^n \frac{\partial}{\partial x_i} (\sqrt{g} g^{ij} \frac{\partial f}{\partial x_j}) = 0.$$

Другими словами, поверхность $F(B)$ является графиком в $\mathbf{R}^{n+k}$ отображения $f: B \rightarrow \mathbf{R}^k$, и система уравнений распадается. Здесь g и g^{ij} имеют прежний смысл, а $g_{ij} = \delta_{ij} + \langle \partial f / \partial x_i, \partial f / \partial x_j \rangle$. Полученная нами система эквивалентна следующей:

$$\sum_{i=1}^n \frac{\partial}{\partial x_i} (\sqrt{g} g^{ij}) = 0, \quad 1 \leq j \leq n; \quad \sum_{i,j=1}^n g^{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} = 0.$$

Сформулируем теперь задачу Дирихле для вектор-функций описанного типа. Пусть задано гладкое отображение φ границы ∂B области B в $\mathbf{R}^k$. Требуется найти отображение $f: B \rightarrow \mathbf{R}^k$ такое, чтобы f было непрерывно на замыкании области B и локально липшицевым в области B и, кроме того, чтобы f удовлетворяло уравнению минимальной поверхности (*). Наконец, требуется, чтобы объем поверхности $F(B)$, т. е. графика отображения f , был конечен (рис. 65). Конечно, можно было бы требовать не гладкости f , а принадлежности f к некоторому более широкому функциональному классу, но мы не будем вдаваться здесь в эти детали. Область B мы предполагаем «хорошей», имеющей, например, кусочно-гладкую границу. Рассмотрим сначала простейший случай, когда коразмерность графика равна единице, т. е. $k=1$. Известно, что в этом случае задача Дирихле для выпуклых областей имеет единственное решение для любых заранее заданных непрерывных граничных условий. Более точно, имеет место:

Предложение 2. При $k=1$ задача Дирихле в выпуклой области $B \subset \mathbf{R}^n$ всегда разрешима для любых непрерывных начальных данных, т. е. для любой непрерывной функции $\varphi: \partial B \rightarrow \mathbf{R}^1$. При этом решение $f: B \rightarrow \mathbf{R}^1$ обладает свойствами:

1) решение единственно, 2) решение вещественно-аналитично, 3) n -мерный график отображения $f: B \rightarrow \mathbf{R}^1$ (т. е. совокупность точек вида $(x, f(x))$ в $\mathbf{R}^{n+1}$, где $x \in B$) является глобально минимальной поверхностью в $\mathbf{R}^{n+1}$, т. е. ее объем наименьший в классе поверхностей с данной границей $(x, \varphi(x))$, где $x \in \partial B \subset \mathbf{R}^n$. Поверхности понимаются здесь в смысле интегральных потоков.

Для случая больших коразмерностей следует особо выделить двумерные выпуклые области $B \subset \mathbf{R}^2$, для которых, как оказывается, задача Дирихле также всегда разрешима.

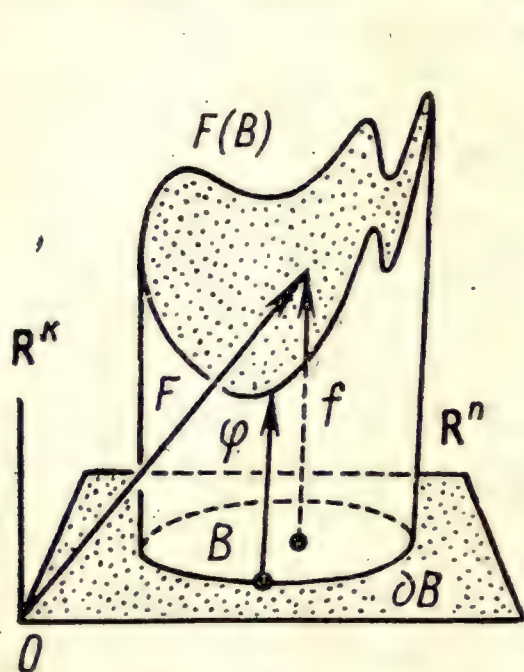


Рис. 65

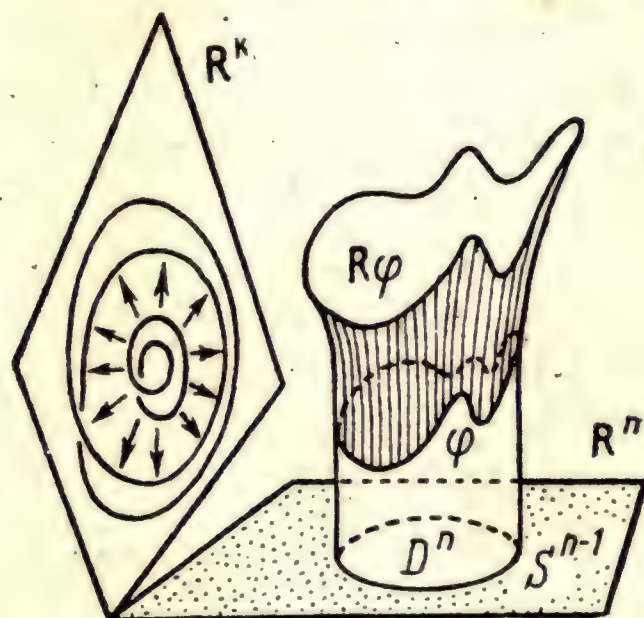


Рис. 66

А именно, известно, что при $n=2$ и при любом $k \geq 1$ для любой выпуклой области $B \subset \mathbf{R}^2$ всегда существует решение задачи Дирихле (непрерывное на замыкании области) для любых непрерывных начальных данных, т. е. для любого непрерывного отображения $\varphi: \partial B \rightarrow \mathbf{R}^k$. Однако, хотя в этом случае теорема существования решения имеет место, тем не менее двумерное решение уже может оказаться не единственным и неустойчивым.

Рассмотрим в качестве B двумерный диск D^2 единичного радиуса на евклидовой плоскости. Тогда существует вещественно-аналитическое отображение $\varphi: \partial D^2 \rightarrow \mathbf{R}^2$ граничной окружности ∂D^2 в $\mathbf{R}^2$ (здесь $n=2$ и $k=1$), что имеются по крайней мере три различных решения задачи Дирихле с одним и тем же граничным условием φ для области D^2 . Более того, одно из этих решений (график отображения $f: D^2 \rightarrow \mathbf{R}^2$) изображается неустойчивой минимальной двумерной поверхностью в четырехмерном евклидовом пространстве, т. е. ее площадь может быть уменьшена сколь угодно малой непрерывной деформацией, неподвижной на границе [30]. Построение примера довольно деликатно, и мы вынуждены опустить его ввиду недостатка места.

Рассмотрим минимальные поверхности больших коразмерностей, т. е. $k \geq 2$. В качестве выпуклой области B в $\mathbf{R}^n$ возьмем стандартный открытый диск D^n единичного радиуса. Ока-

зывается, при $n \geq 4$ задача Дирихле может не иметь решения. Существует отображение $f: \partial D^n \rightarrow \mathbb{R}^k$ для некоторого k , где $2 \leq k \leq n-2$, такое, что задача Дирихле для уравнения минимальной поверхности в шаре D^n с граничными условиями f на сфере $S^{n-1} = \partial D^n$ не имеет никакого липшицева решения F , такого, что $F|_{S^{n-1}} = f$. Любопытно, что построение таких отображений опирается на топологические соображения [30].

Предложение 3. Пусть $n > k \geq 2$, $S^{n-1} = \partial D^n \subset \mathbb{R}^n$, $S^{k-1} \subset \mathbb{R}^k$, где S^{n-1} и S^{k-1} — стандартные сферы единичного радиуса. Пусть $\varphi: S^{n-1} \rightarrow S^{k-1}$ — любое отображение класса C^2 (т. е. имеющее непрерывные производные до второго порядка включительно), определяющее нетривиальный элемент гомотопической группы $\pi_{n-1}(S^{k-1})$, т. е. негомотопное нулю. Тогда существует число R_φ такое, что для каждого $R \geq R_\varphi$ не существует решения задачи Дирихле в шаре D^n для уравнения минимальной поверхности коразмерности k с граничной функцией $\varphi_R = R \cdot \varphi$ (рис. 66).

С геометрической точки зрения происходит следующее. Мы рассматриваем график отображения, реализующего ненулевой элемент гомотопической группы, а затем (рис. 66) подвергаем график простому преобразованию — растягиваем с коэффициентом подобия R . Центром подобия является начало координат в $\mathbb{R}^k$. Наряду с растяжением граничных данных можно рассмотреть обратный процесс сжатия граничной функции путем умножения ее на уменьшающийся параметр. Этот процесс интересен тем, что задача Дирихле всегда разрешима для выпуклой области (и для любой коразмерности) для достаточно малых граничных данных. Другими словами, если функция φ мала по абсолютной величине на ∂B , то существует минимальная поверхность, близкая к B и совпадающая на границе B с графиком φ . Отсюда следует, что, умножая граничную функцию на уменьшающийся параметр r , мы в конце концов найдем требуемую минимальную поверхность. Иначе говоря, существует число $r_\varphi > 0$ такое, что для всех $r < r_\varphi$ задача Дирихле разрешима в области B для граничных данных $\varphi_r = r \cdot \varphi$.

В качестве примера рассмотрим отображение Хопфа, подробно изученное нами в гл. 1. Это расслоение задается отображением $p: S^3 \rightarrow S^2$, которое записывается в координатах z, w так: $p(z, w) = (|z|^2 - |w|^2, 2z\bar{w})$. Здесь z, w — стандартные эрмитовы координаты в $\mathbb{C}^2$, сфера S^3 единичного радиуса задается обычным образом: $|z|^2 + |w|^2 = 1$, а сфера S^2 реализована как единичная сфера в $\mathbb{R}^3 = \mathbb{R} \times \mathbb{C}$. В самом деле, легко убедиться непосредственным вычислением, что $(|z|^2 - |w|^2)^2 + (\operatorname{Re} 2z\bar{w})^2 + (\operatorname{Im} 2z\bar{w})^2 = 1$, т. е. что вектор $p(z, w)$ пробегает двумерную сферу в $\mathbb{R}^3$. Как известно (см., например, [4]), отображение p негомотопно нулю, т. е. определяет нетривиальный элемент группы $\pi_3(S^2) = \mathbb{Z}$. Следовательно, применяя предложение 3, мы получаем, что при достаточно больших R задача Дирихле на четырехмерном шаре с граничными условиями

$R \cdot \varphi$ не имеет решения в виде минимального графика некоторого отображения $F: D^4 \rightarrow \mathbf{R}^3$. Прямое вычисление (которое мы здесь опускаем) показывает, что в качестве граничного значения R_φ можно взять 4,2 при единичных радиусах сфер. Следовательно, при всех $R > 4,2$ задача Дирихле с граничной функцией $R \cdot \varphi$ не имеет решения [30].

Следует помнить, что отсутствие решения задачи Дирихле при указанных ограничениях отнюдь не препятствует наличию четырехмерной минимальной поверхности, имеющей в качестве граничного «контура» трехмерную сферу, вложенную в $\mathbf{R}^7$ как график отображения $\varphi: S^3 \rightarrow S^2 \subset \mathbf{R}^3$. Более того, как следует из общих теорем существования, для любого непрерывного отображения φ (в том числе и указанного вида) всегда существует глобально минимальная поверхность, затачивающая этот «контур» в $\mathbf{R}^7$. Из предложения 3 вытекает, что эта поверхность не может быть представлена в виде графика какого-либо отображения $F: D^4 \rightarrow \mathbf{R}^3$.

Существование решения задачи Дирихле с граничными данными $R \cdot \varphi$ при достаточно малых R и, напротив, отсутствие решения, когда параметр R велик, указывает на то, что имеется некоторое граничное критическое значение параметра R_0 , отделяющее друг от друга эти две зоны качественно различного поведения многомерной вариационной задачи. Интуитивно ясно, что при $R = R_0$ должно существовать какое-то сингулярное решение задачи Дирихле для граничной функции $R_0 \times \varphi$. В частном случае отображения Хопфа наличие значительных симметрий у этого расслоения позволяет предположить, что сингулярное решение должно изображаться конусом, границей которого должно быть вложение сферы S^3 в $\mathbf{R}^7$, определяемое заданием граничной функции $R_0 \cdot \varphi$. Оказывается, это предположение оправдывается. Таким образом, в этой вариационной задаче снова возникают минимальные конусы, с которыми мы неоднократно сталкивались и раньше.

Рассмотрим гладкое вложение сферы S^3 в сферу S^6 , задаваемое так: $i_\alpha(x) = (\alpha x, \sqrt{1 - \alpha^2} \cdot p(x))$, где $p: S^3 \rightarrow S^2$ — отображение Хопфа. Здесь $0 < \alpha \leq 1$, x — точка на единичной сфере S^3 , пространство $\mathbf{R}^7$ представлено в виде прямого произведения $\mathbf{R}^4 \times \mathbf{R}^3$, т. е. $\mathbf{R}^7 = \mathbf{C}^2 \times \mathbf{R}^3$, $\alpha x \in \mathbf{C}^2$, $\sqrt{1 - \alpha^2} p(x) \in \mathbf{R}^3$. Так как $|x| = 1$ для $x \in S^3$ и $|p(x)| = 1$ для $p(x) \in S^2$, то $|\alpha(x)|^2 + |\sqrt{1 - \alpha^2} p(x)|^2 = 1$, т. е. $i_\alpha(x) \in S^6 \subset \mathbf{R}^7$. Рассмотрим гладкое действие группы $SU(2)$ на $\mathbf{R}^7 = \mathbf{C}^2 \times \mathbf{R}^3$. Для этого рассмотрим диагональное вложение группы $SU(2)$ в группу $SU(2) \times SO(3)$, где каждый из сомножителей действует естественным образом, согласованным с разложением $\mathbf{R}^7$ в прямое произведение. Напомним, что группа $SU(2)$ действует на $\mathbf{C}^2$ как группа специальных унитарных преобразований, а на $\mathbf{R}^3$ действует как группа $SO(3) = SU(2)/\mathbf{Z}_2$, где $\mathbf{Z}_2$ — центр. Очевидно, что отображение Хопфа $p: S^3 \rightarrow S^2$ эквивариантно отно-

сительно этих действий группы $SU(2)$. Отсюда, в частности, следует, что график отображения является $SU(2)$ -инвариантной поверхностью в $\mathbb{R}^7$. Так как действие группы $SU(2)$ транзитивно на сфере S^3 , то график также является орбитой любой своей точки при указанном действии группы $SU(2)$.

Итак, при каждом α (где $0 < \alpha \leq 1$) сфера $i_\alpha S^3$, гладко вложенная в сферу S^6 , является орбитой гладкого действия группы $SU(2)$ на S^6 . Легко убедиться в том, что это — орбиты общего положения, главные орбиты (в смысле п. 7 настоящего параграфа). Как мы знаем, локально минимальные орбиты взаимно-однозначно соответствуют критическим точкам функции объема. В частности, орбиты максимального объема являются минимальными подмногообразиями в сфере S^6 . Обозначим через $v(\alpha)$ объем орбиты $i_\alpha S^3$. Прямое вычисление показывает, что $v(\alpha) = 2\pi^2 \alpha (4 - 3\alpha^2)^2$. Эта функция от α достигает своего максимума на отрезке $[0, 1]$ при $\alpha = 2/3$. Следовательно, сфера $i_{2/3} S^3$, гладко вложенная в сферу S^6 при помощи отображения $x \rightarrow \left(\frac{2}{3} x, \frac{\sqrt{5}}{3} p(x) \right)$, где $x \in S^3$ и $|x| = 1$, является локально-минимальным подмногообразием в S^6 . Здесь $\frac{2}{3} x \in \mathbb{R}^4$, $\frac{\sqrt{5}}{3} p(x) \in \mathbb{R}^3$.

Напомним, что конус CA с вершиной в точке O над гладким подмногообразием A^q в сфере S^m является локально-минимальным в том и только в том случае, когда подмногообразие A^q является локально-минимальным в сфере S^m . Следовательно, четырехмерный конус $C(i_{2/3} S^3)$ с вершиной в точке O , имеющий границей подмногообразие $i_{2/3} S^3$ в S^6 , является минимальной поверхностью в $\mathbb{R}^7$, имеющей в точке O особенность. С другой стороны, этот конус, очевидно, определяет некоторое липшицево отображение $f_0: D^4 \rightarrow \mathbb{R}^3$, являющееся решением задачи Дирихле для шара D^4 с граничным условием $R_0 \cdot \varphi$ (см. выше). Имеет место утверждение: липшицево отображение $f_0: \mathbb{R}^4 \rightarrow \mathbb{R}^3$, задаваемое формулой $f_0(x) = \frac{\sqrt{5}}{2} |x| p\left(\frac{x}{|x|}\right)$ при $x \neq 0$, где p является отображением Хопфа, является решением уравнения минимальной поверхности (*) и, в частности, сингулярным решением задачи Дирихле. Как мы видим, существуют липшицевы решения системы (*), не принадлежащие классу гладкости C^1 .

Аналогичные примеры конструируются и при помощи двух других расслоений Хопфа, а именно $p': S^7 \rightarrow S^4$ и $p'': S^{15} \rightarrow S^8$. Соответствующие функции объема орбит имеют вид:

$$v'(\alpha) = \frac{\pi^4}{3} \alpha^3 (4 - 3\alpha^2)^2; \quad v''(\alpha) = \frac{2\pi^8}{7!} \alpha^7 (4 - 3\alpha^2)^4.$$

15. Пример гладкой замкнутой незаузленной кривой в $\mathbb{R}^3$,

которая ограничивает только минимальные поверхности большого рода.

Как мы знаем, на любую замкнутую гладкую кривую A в $\mathbb{R}^3$ натягивается минимальная поверхность, имеющая кривую A своей границей в смысле тех определений поверхности и ее границы, которые обсуждались выше. Если кривая заузлена, т. е. образует нетривиальный узел (не эквивалентный стандартному плоскому вложению окружности в плоскость), то минимальная поверхность имеет, вообще говоря, особенности. При этом минимум площади может достигаться на какой-либо двумерной поверхности, отличной от диска, т. е. получающейся из диска приклейкой некоторого количества ручек. Что происходит в случае незаузленной кривой? Альмгрен и Тёрстен по-

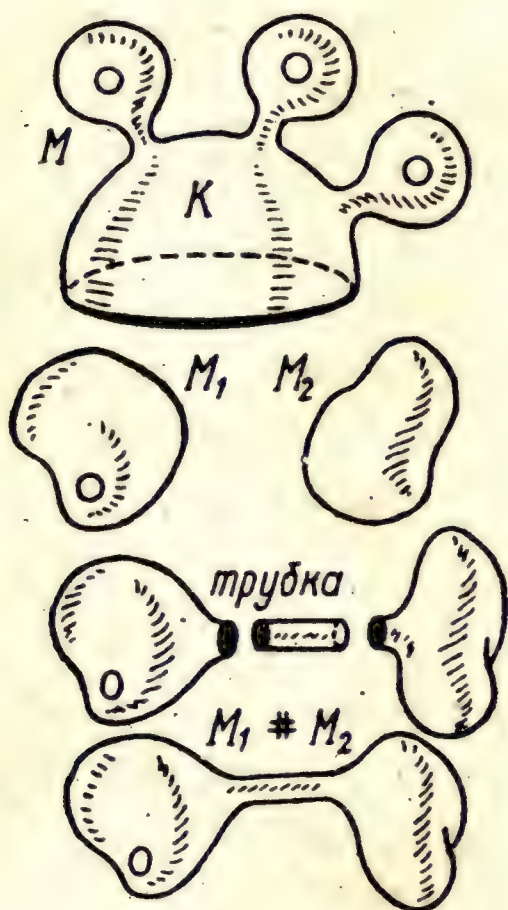


Рис. 67

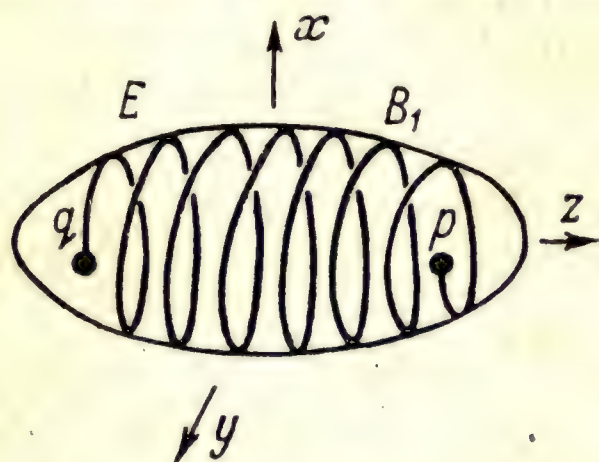


Рис. 68

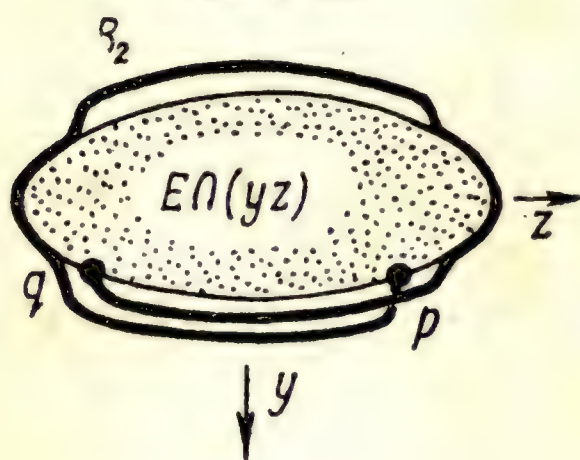


Рис. 69

строили следующий пример кривой, поучительный во многих отношениях. Для любого положительного целого числа g конструируется гладкая замкнутая несамопересекающаяся незаузленная кривая A в $\mathbb{R}^3$ такая, что если M^2 — некоторая поверхность, имеющая кривую A своей границей и целиком расположенная внутри выпуклой оболочки кривой, то поверхность M имеет по крайней мере g ручек. В частности, она не может быть диском. При этом поверхность M предполагается гладко вложенной в $\mathbb{R}^3$. Напомним, что выпуклой оболочкой какого-либо подмножества H в $\mathbb{R}^n$ называется множество точек, замкнутое всевозможными отрезками вида AB , где точки A и B — концы отрезка — принадлежат подмножеству H . Мы выделили условие включения поверхности M в выпуклую оболочку кривой A по той причине, что любая минимальная поверхность, заклеивающая в $\mathbb{R}^n$ некоторый «контур» A^{k-1} , обязательно со-

держится в выпуклой оболочке H этого контура. Будем говорить, что поверхность M имеет по крайней мере g ручек, если она гомеоморфна следующей связной сумме: $K \# T^2 \# \dots \# T^2$ (g раз), где K — некоторое двумерное многообразие, T^2 — тор, а $\#$ — связная сумма многообразий (рис. 67).

Наиболее важным шагом в построении указанной кривой является конструкция ориентированной гладкой замкнутой незаузленной кривой A такой, что если M^2 — любая ориентированная поверхность, расположенная внутри выпуклой оболочки кривой и имеющая кривую своей границей, то M имеет по крайней мере три ручки. Из конструкции легко усмотреть, как построить кривую с аналогичным свойством для любого наперед заданного числа ручек g . Условие ориентируемости существенно, поскольку если отказаться от него, то на эту же кривую можно натянуть лист Мебиуса, имеющий ее своей границей и целиком расположенный внутри выпуклой оболочки кривой. Рассмотрим эллипсоид E , задаваемый в $\mathbf{R}^3(x, y, z)$ уравнением $x^2 + y^2 + \frac{z^2}{16} = 1$, и построим на нем спиральную

траекторию B_1 (рис. 68). Кривая обходит вокруг эллипсоида семь раз и кончается в точках p и q . Увеличивая число оборотов кривой B_1 вокруг оси эллипсоида (семь, восемь, и т. д.),

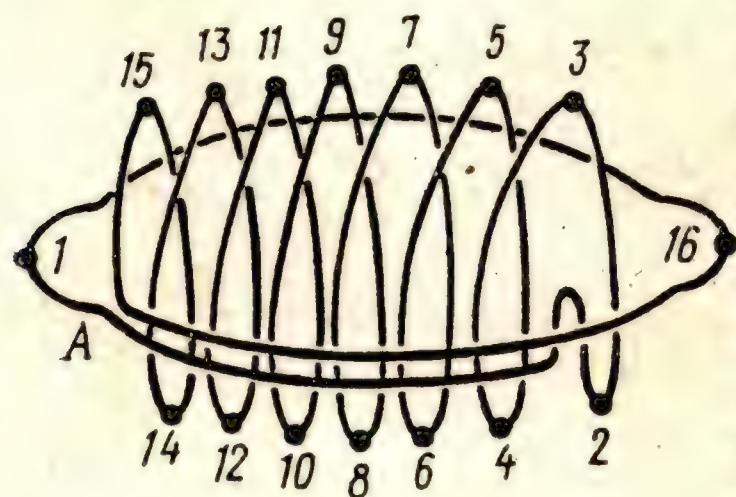


Рис. 70

мы, оказывается, будем увеличивать число ручек минимальной поверхности, заклеивающей конструируемую кривую A . Кривая B_2 , замыкающая кривую B_1 , показана на рис. 69. Так как при построении искомой кривой следует соблюдать некоторые метрические ограничения, то поясним конструкцию B_2 более подробно. Кривая начинается в точке q и продолжается в плоскости (yz)

или непосредственно около нее, двигаясь против часовой стрелки до тех пор, пока не дойдет до точки p . На кривой B_2 выделяются следующие участки: 1) при $-3 < z < 3,5$ и при $y > 0$ кривая начинается в точке q , движется в плоскости (yz) вне эллипсоида E , однако, в непосредственной близости от него, оставаясь приблизительно на одном и том же расстоянии от эллипсоида и двигаясь параллельно кривой пересечения эллипсоида с плоскостью (yz) ; 2) при $z \geq 3,5$ кривая B_2 совпадает с пересечением $E \cap (yz)$; 3) при $-3,5 < z < 3,5$ кривая B_2 движется в плоскости (yz) вне эллипсоида, недалеко от него и параллельно линии его пересечения с плоскостью (yz) ; 4) при $z \leq -3,5$ кривая B_2 снова совпадает с $E \cap (yz)$; 5) при $-3,5 < z < -2,5$ кривая B_2 движется в плоскости (yz) вне эллипсоида, параллельно его пересечению с плоскостью, однако B_2 движется по-прежнему без самопересечений; 6) наконец, при $2,5 < z < 3$

кривая B_2 заканчивается в точке p . При этом, чтобы избежать самопересечения, она слегка опускается под плоскость (yz) , т. е. приобретает незначительное отрицательное приращение по координате x . Объединяя кривые B_1 и B_2 , мы получаем окончательную кривую A (рис. 70). Легко убедиться, что она незаузлена, т. е. гладкой деформацией превращается без самопересечений в стандартную окружность. Пусть $\bar{A}$ — выпуклая оболочка кривой A . Ясно, что если рассмотреть часть кривой A , состоящую из точек (x, y, z) , для которых координата y достаточно мала, то эти точки расположены на границе $\bar{A}$. Пусть M — гладкая ориентированная поверхность, целиком содержащаяся в $\bar{A}$ и имеющая A своей границей. Мы хотим доказать, что количество ручек у поверхности не меньше трех. Выберем достаточно малое t такое, чтобы плоскость Π , задаваемая уравнением $y=t$, пересекалась трансверсально с поверхностью M и кривой A . Из конструкции граничной кривой следует, что ее пересечение с Π состоит из 16 изолированных точек, занумерованных числами от 1 до 16 на рис. 70. Нумерация точек соответствует относительному упорядочиванию, определенному ориентацией кривой A . Пометим точки с нечетными номерами знаком «+», а точки с четными номерами — знаком «—». Обозначения согласуются с тем, что кривая A пересекает Π в направлении увеличения координаты y в точках, помеченных «+», и, напротив, кривая пересекает Π в направлении уменьшения y в точках, помеченных «—». Все 16 точек пересечения кривой с плоскостью расположены на границе выпуклой оболочки кривой. Так как точки типа «+» и «—» расположены на крае многообразия M , то пересечение $M \cap \Pi$ состоит из восьми ориентированных гладких дуг, соединяющих 8 положительных точек с восемью отрицательными точками, и, быть может, кроме этих 8 дуг есть еще некоторое число замкнутых кривых, гомотомичных окружности. Эти последние кривые нас не будут интересовать. Каждая из восьми дуг начинается в некоторой положительной (нечетной) точке и заканчивается в некоторой отрицательной (четной) точке. Обозначим эти восемь дуг символами $\gamma_1, \gamma_3, \gamma_5, \dots, \gamma_{15}$, указывая номер положительной точки, являющейся одним из концов дуги. Выясним, в какие отрицательные точки идут эти дуги, т. е. какие четные номера следует приписать другим концам этих дуг. Рассмотрим, например, дугу γ_1 , выходящую из точки 1. Из рис. 70 видно, что единственной точкой, в которой может закончиться эта дуга, является точка 14. В самом деле, если бы дуга кончилась в какой-либо другой точке (например, в точке 12), то точка 14 была бы «отрезана» этой дугой от всех остальных положительных (нечетных) точек и никакая дуга γ_i , отличная от γ_1 , не смогла бы закончиться в точке 14. Здесь существенно использовано то, что поверхность M расположена целиком внутри выпуклой оболочки кривой A . В силу трансверсальности пересечения M с Π дуги γ_i попарно не пересекаются, начинаются и

заканчиваются на границе M , т. е. на A . Опуская перебор всех возможных вариантов, получаем следующую таблицу соответствия начальных и конечных точек восьми дуг: γ_1 заканчивается в точке 14, γ_3 — в точке 16, γ_5 — в точке 2, γ_7 — в точке 4, γ_9 — в точке 6, γ_{11} — в точке 8, γ_{13} — в точке 10, γ_{15} — в точке 12.

Оценим снизу количество ручек поверхности M . Для этого удобно немного расширить M . Заменим ее на гомеоморфную поверхность M^* , получающуюся из M ее незначительным продолжением за край A . Другими словами, приклеим к краю A .

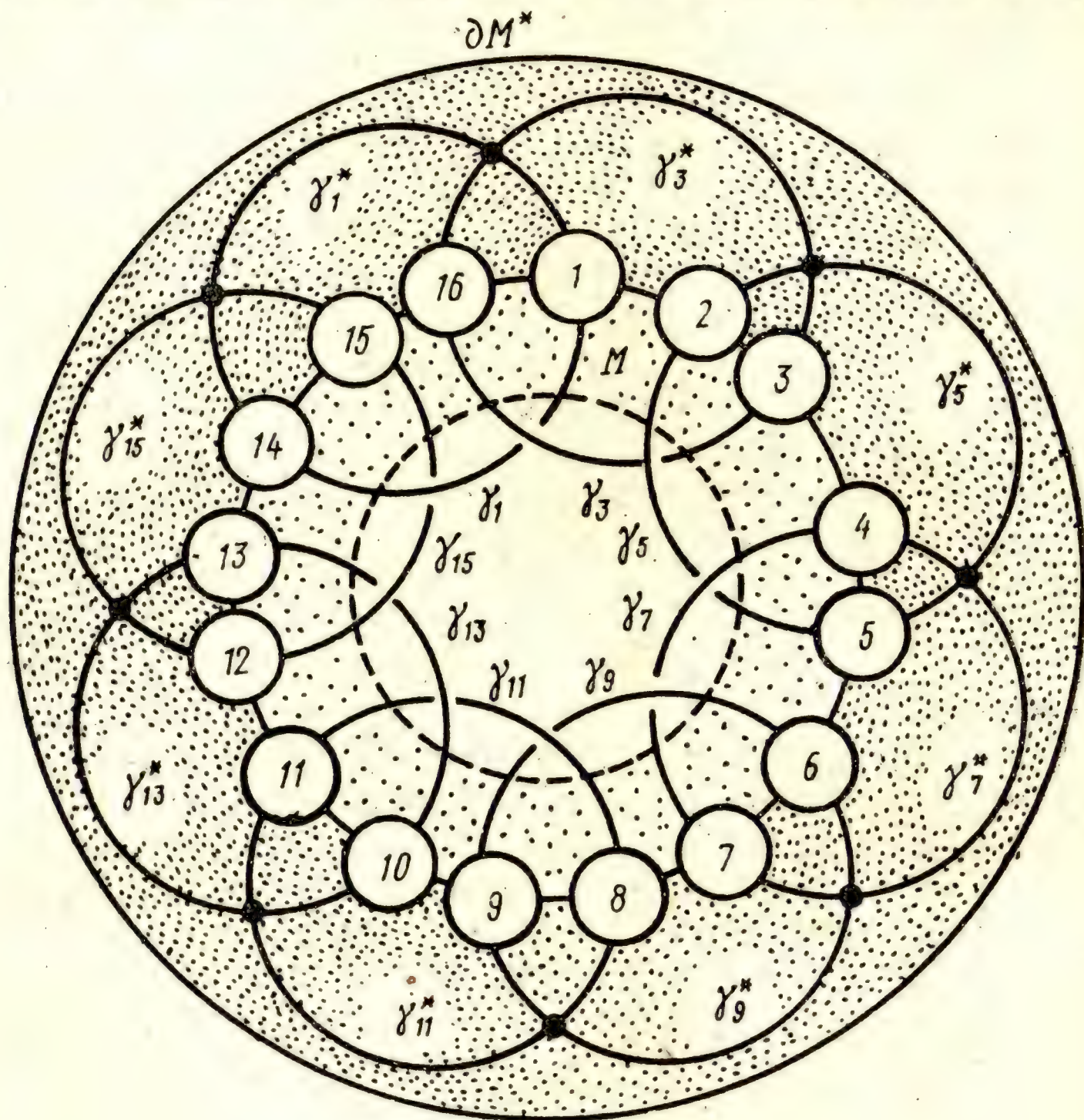


Рис. 71

цилиндр $S^1 \times I$ по одному из его краев. Замкнем дуги γ_i , добавив к ним отрезки, расположенные в добавленном нами кольце (рис. 71). В результате каждая дуга γ_i продолжается до замкнутой гладкой кривой γ_i^* . Каждая из них определяет некоторый одномерный цикл $[\gamma_i^*]$, т. е. элемент группы гомологий $H_1(M, \mathbb{Z}_2)$, где в качестве группы коэффициентов взята группа $\mathbb{Z}_2$. Вычислим матрицу пересечений полученных одномерных циклов $[\gamma_i^*]$. Это квадратная матрица размера 8×8 с коэффициентами 0 или 1. Нумеруя столбцы и строки матрицы циклами $[\gamma_1^*]$, $[\gamma_3^*]$, ..., $[\gamma_{15}^*]$, получаем следующую матрицу, эле-

менты которой однозначно восстанавливаются из рис. 71:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Ясно, что ранг матрицы не меньше, чем 6, так как строки со 2-й по 7-ю независимы. Операция взятия пересечения с каким-либо фиксированным одномерным циклом определяет линейную функцию над $\mathbb{Z}_2$ на векторном пространстве $H_1(M^*, \mathbb{Z}_2)$. Так как мы обнаружили по крайней мере 6 линейно-независи-

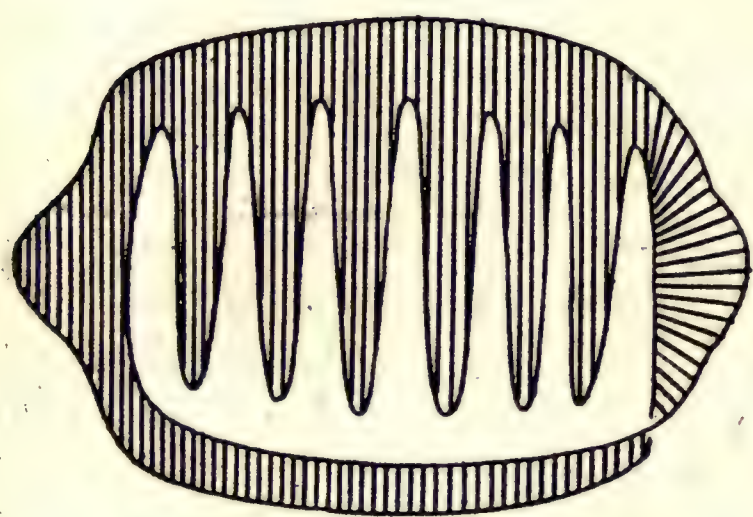


Рис. 72

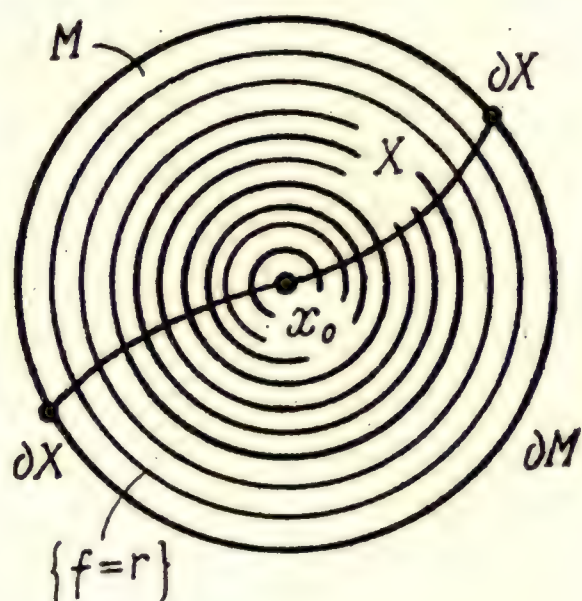


Рис. 73(a)

мых функционалов на $H_1(M^*, \mathbb{Z}_2)$, то размерность $H_1(M, \mathbb{Z}_2)$ не меньше шести, ибо $H_1(M, \mathbb{Z}_2) = H_1(M^*, \mathbb{Z}_2)$. Поскольку каждая ручка дает вклад в одномерную группу гомологий в виде двумерного подпространства, то, чтобы набрать размерность 6, требуется по крайней мере 3 ручки, что и требовалось.

Отказ от ориентируемости сразу приводит к появлению гладкой поверхности, имеющей A границей и гомеоморфной листу Мебиуса. Мы фактически знакомы с этой реализацией листа Мебиуса по рис. 48 настоящей главы. На рис. 72 изображена ортогональная проекция кривой A на плоскость (yz) . Каждый отрезок, показанный на рис. 72, является проекцией отрезка из $\mathbb{R}^3$, соединяющего две точки граничной кривой. Лист Мебиуса получается как объединение всех этих отрезков, к которому добавлен еще один отрезок, проектирующийся в точку скрещивания двух ветвей кривой A (рис. 72).

Развивая описанную конструкцию, можно построить пример гладкой замкнутой несамопересекающейся незаузленной кривой в $\mathbb{R}^3$ такой, что если M^2 — некоторая поверхность (уже не обязательно ориентируемая) с границей A , целиком лежащая

в ее выпуклой оболочке, то M имеет по крайней мере три ручки. Детали построения мы опускаем.

Существование кривой с указанными свойствами оказывается полезным для более глубокого понимания геометрии минимальных поверхностей.

Как мы отмечали, существует много вариантов определения границы поверхности. Так, например, известно, что каждая гладкая замкнутая ориентированная кривая A в $\mathbb{R}^3$ является границей двумерного интегрального потока M , масса которого наименьшая по сравнению с массами всех интегральных потоков с той же границей. Носитель этого минимального потока (за исключением граничных точек) является гладким двумерным минимальным подмногообразием в $\mathbb{R}^3$. Используя проведенные выше рассуждения, можно показать, что носитель должен содержать по крайней мере три ручки. Другой подход заключается в нахождении отображения фиксированной двумерной области с некоторым количеством ручек в $\mathbb{R}^3$ такого, чтобы граница области отображалась на кривую A и чтобы площадь образа была наименьшей возможной. Известно, что если область является диском, то отображение с наименьшей площадью существует для любой гладкой замкнутой связной несамопересекающейся кривой A в $\mathbb{R}^3$. При этом во внутренних точках минимальное отображение является погружением. Если же в качестве границы взять кривую A , сконструированную выше, то не существует минимального вложения гладкой двумерной области в $\mathbb{R}^3$ с границей A , если область имеет мало ручек. В частности, не существует минимального вложения диска с границей A .

§ 3. Геометрия экстремалей функционала объема и функционала Дирихле

1. Оценка снизу объемов минимальных поверхностей.

Пусть X^k — глобально минимальная поверхность нетривиального топологического типа в многообразии M^n , например, X реализует некоторый цикл из группы $H_k(M, G)$ или некоторый класс бордизмов из группы $\Omega_k(M)$. При изучении свойств таких поверхностей часто возникает важный вопрос об оценке снизу объема поверхности. Другими словами, каков наименьший объем k -мерных циклов многообразия M , т. е. наименьший объем носителей этих циклов? Точные оценки такого рода особенно важны для конкретного предъявления глобально минимальных поверхностей, поскольку для доказательства минимальности конкретной поверхности следует убедиться, что ее объем равен нижней грани объемов поверхностей того же топологического типа. А для этого нужно уметь эффективно и в явном виде вычислять эту нижнюю грань. Нетривиальность задачи связана с тем, что требуется исследование больших ва-

риаций, а не бесконечно малых, как это обычно делается при доказательстве локальной минимальности какой-либо поверхности. При этом нижняя оценка на объемы должна быть универсальной и связывать метрические и топологические свойства поверхностей. Оказывается, существует эффективный способ вычислять наименьшие объемы замкнутых поверхностей нетривиального топологического типа в терминах метрики объемлющего многообразия [16, 27] (А. Т. Фоменко).

Пусть M^n — гладкое компактное ориентируемое связное риманово многообразие с краем ∂M . Пусть $f(x)$ — гладкая функция на M такая, что она имеет единственную точку x_0 минимума (который, следовательно, является абсолютным), изменяется от 0 до 1 на M , достигает максимума только на крае ∂M и равна 1 во всех точках края. Пусть на ∂M нет никаких критических точек функции f и пусть f — функция Морса на открытом многообразии $M \setminus (\partial M \cup x_0)$ и имеет лишь конечное число критических точек. Ясно, что среди этих точек нет локальных минимумов или максимумов. Рассмотрим поверхность уровня $\{f=r\}$, т. е. множество точек x , где $f(x)=r$, $0 \leq r \leq 1$. Имеем: $\{f=0\}=x_0$, $\{f=1\}=\partial M$. Рассмотрим гладкое векторное поле $v = \text{grad } f$ на M . Пусть X — произвольное k -мерное измеримое компактное подмножество в M . Рассмотрим «границу» этого компакта X , т. е. множество $\partial X = X \cap \partial M$ (рис. 73(а)).

Напомним понятие стандартной функции $\Psi(x)$ сферической плотности подмножества X в M . Рассмотрим точку x и пусть B_ε — n -мерный шар геодезического радиуса ε с центром в точке x . Рассмотрим его пересечение с X и подсчитаем k -мерный объем (меру) этого пересечения $\text{vol}_k X \cap B_\varepsilon$. Вычислим предел $\lim_{\varepsilon \rightarrow 0} \frac{\text{vol}_k X \cap B_\varepsilon}{p_\varepsilon} = \Psi(x)$, где $p_\varepsilon = (k\text{-мерный объем стандартного } k\text{-мерного евклидова шара радиуса } \varepsilon)$. Функция $\Psi(x)$ называется плотностью X в точке x . Если $x \notin X$, то $\Psi(x) = 0$. Если $x \in X$ и x является регулярной точкой, т. е. подмножество X — гладкое подмногообразие в окрестности этой точки, то $\Psi(x) = 1$. Это следует из того, что при малом ε пересечение $X \cap B_\varepsilon$ мало отличается от плоского k -мерного диска радиуса ε . Если X имеет в окрестности точки x вид острого «тонкого уса» (рис. 73(б)), то k -мерный объем пересечения X с шаром B_ε мал и, следовательно, $\Psi(x) < 1$. Если, напротив, точка $x \in X$ является особой точкой, например, если в ней сходится много k -мерных листов поверхности X (как это имеет место в случае двумерных мыльных пленок), то $\Psi(x) > 1$. Итак, функция плотности достаточно чутко реагирует на локальные свойства X .

Рассмотрим вариационную задачу в классе поверхностей, заклеивающих фиксированный контур A , где A — $(k-1)$ -мерное измеримое компактное подмножество в ∂M такое, что $H_{k-1}(A, G) = 0$. Введем класс поверхностей, для которого будем

решать задачу о существовании абсолютного минимума. Рассмотрим класс A^* всех компактов X с границей $A \subset \partial M$ таких, что: 1) $\dim X = k$, компакт X измерим; 2) X не содержит «тонких усов», т. е. для любой точки $x \in X \setminus \partial X$ имеем: $\Psi(x) \geq 1$; 3) граница X содержит A , т. е. $A \subset \partial X$; 4) при вложении A в X аннулируется вся группа $(k-1)$ -мерных гомологий A , т. е. гомоморфизм $i_*: H_{k-1}(A, G) \rightarrow H_{k-1}(X, G)$, индуцированный вложением $i: A \rightarrow X$, тривиален.

Здесь H_{k-1} — группы спектральных гомологий (см. выше). Если d — нижняя грань объемов всех поверхностей из класса A^* , то существует глобально минимальная поверхность $X_0 \in A^*$ такая, что $\text{vol}_k X_0 = d$. Все эти и нижеследующие результаты настоящего пункта верны и для когомологий (в том числе и обобщенных), но мы не будем на этом останавливаться [16,

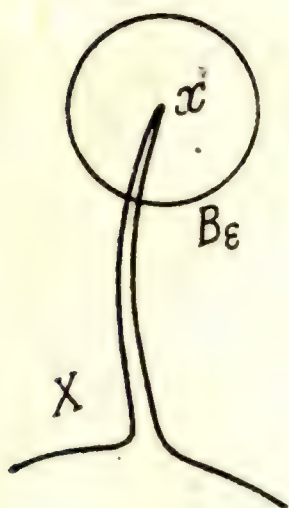


Рис. 73(б)

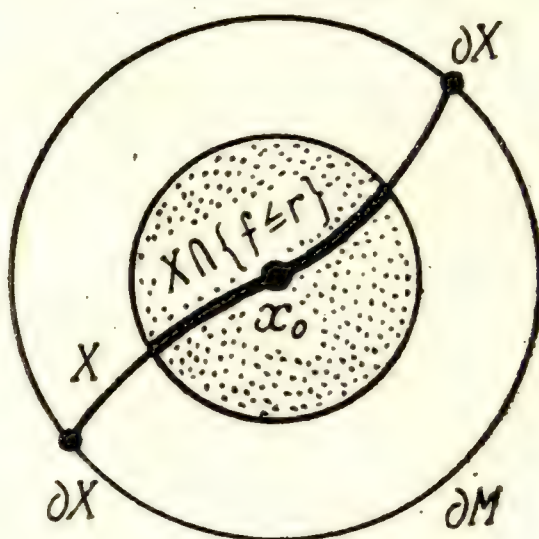


Рис. 74



Рис. 75

27]. Для поверхности X_0 также имеем: $\Psi_k(x) \geq 1$ при любом $x \in X_0 \setminus \partial X_0$. Будем считать, что X_0 проходит через точку x_0 , т. е. точку абсолютного минимума функции f . Построим важную функцию, которую назовем функцией объема поверхности: $\psi(r) = \text{vol}_k X_0 \cap \{f \leq r\}$. Через $\{f \leq r\}$ обозначено n -мерное подмножество в M , состоящее из точек x , где $f(x) \leq r$ (рис. 74). Для разных поверхностей X_0 функция объема будет, вообще говоря, различной. Функция объема определена на отрезке $[0, 1]$ и, очевидно, не убывает с ростом r от 0 до 1. Задачу, описанную в начале пункта, можно теперь сформулировать так: дать точную оценку снизу на функцию ψ в терминах метрики M независимо от топологического типа минимальной поверхности X_0 . Точность понимается в том смысле, что искомая оценка должна превращаться в точное равенство для достаточно богатых серий конкретных троек (M, f, X) .

Ясно, что $\psi(1) = \text{vol}_k X$, так как $X \cap \{f \leq 1\} = X \cap M = X$. Итак, зная нижнюю оценку на функцию $\psi(r)$, мы сможем оценить снизу и число $\psi(1)$, т. е. объем минимальной поверхности. При

этом мы можем менять границу A в ∂M , следя лишь за тем, чтобы группа $H_{k-1}(A, G)$ была нетривиальной и аннулировалась гомоморфизмом i_* .

2. Коэффициент деформации векторного поля.

Введем понятие коэффициента деформации векторного поля $\text{grad } f$. В силу свойств функции f почти все интегральные траектории поля, начинаясь в x_0 , заканчиваются на крае ∂M . Будем считать, что точка $x \in M$ является точкой общего положения, если через нее проходит интегральная траектория, начинающаяся в x_0 и кончающаяся на ∂M . Рассмотрим в многообразии M в точке x малый $(k-1)$ -мерный шар D_ϵ радиуса ϵ , ортогональный вектору $\text{grad } f(x)$ (рис. 75). Рассмотрим множество всех интегральных траекторий, которые начинаются в точке x_0 и заканчиваются в точках этого шара. Обозначим множество точек, заметаемое этими траекториями, через CD_ϵ . С геометрической точки зрения это k -мерное множество является «конусом», имеющим вершину в точке x_0 и в качестве основания — шар D_ϵ . «Конус» получится, если мы начнем деформировать шар D_ϵ , двигая его вдоль интегральных траекторий поля $\text{grad } f$ по направлению к точке x_0 . Шар при своем движении замечает множество CD_ϵ . Если точка x — общего положения, то конус является гладким k -мерным подмногообразием с краем в M . Если точка x произвольна, то конус определяется аналогично, но в этом случае он будет, вообще говоря, не подмногообразием, а клеточным комплексом.

Вычислим число $\lambda(x) = \lim_{\epsilon \rightarrow 0} \frac{\text{vol}_k CD_\epsilon}{\text{vol}_{k-1} D_\epsilon}$. Другими словами, вы-

числяем отношение объема конуса к объему его основания и переходим к пределу, когда основание стягивается в точку. Ясно, что число $\lambda(x)$ зависит от выбора $(k-1)$ -мерного диска D_ϵ , ортогонального вектору $\text{grad } f(x)$. Рассмотрим множество всех таких $(k-1)$ -мерных дисков, и пусть $\kappa_k(x) = \sup \lambda(x)$, где $\sup$ берется по всем дискам D_ϵ , ортогональным вектору $\text{grad } f(x)$. Функция $\kappa_k(x)$ является гладкой на M . Назовем ее k -мерным коэффициентом деформации векторного поля $\text{grad } f$. Для разных x коэффициент будет, вообще говоря, различным, поскольку он тесно связан с тензором римановой кривизны многообразия M . Во многих случаях $\kappa_k(x)$ легко вычисляется. Рассмотрим пример. Пусть M — шар радиуса R в $\mathbf{R}^n$, пусть $f(x) = |x|$ — расстояние от точки $0 = x_0$ до точки x , $v = \text{grad } f$. Тогда $\kappa_k(x) = r/k$, где $r = |x| = f(x)$.

3. Замкнутые поверхности нетривиального топологического типа и наименьшего объема.

Введенный нами коэффициент деформации поля играет важную роль в вычислении наименьших объемов глобально минимальных поверхностей (А. Т. Фоменко).

Теорема 1. (см. [16]). Пусть f — гладкая функция на M^n такая, что $0 \leq f \leq 1$, $f(x_0) = 0$, $f(\partial M) = 1$; f является функ-

цией Морса на $M \setminus (x_0 \cup \partial M)$, имеющей лишь седловые критические точки, индексы которых не превосходят числа $k-2$, где k — целое число и $k < n$. Пусть X_0^k — глобально минимальная поверхность описанного класса, проходящая через x_0 . Тогда для функции объема этой поверхности выполнено неравенство: $\psi(r) \geq q(r) \cdot l$, где постоянная $l = \lim_{a \rightarrow 0} \frac{\psi(a)}{q(a)}$ не зависит от параметра r и определяется только структурой минимальной поверхности X_0 около точки x_0 , а функция $q(r)$ имеет вид:

$$q(r) = \exp \int \left[\max_{x \in \{f=r\}} (\kappa_k(x) \cdot |\operatorname{grad} f(x)|) \right]^{-1} dr.$$

В частности, $\psi(1) = \operatorname{vol}_k X_0 \geq l \cdot q(1)$. Итак, поведение функции объема $\psi(r)$ на поверхности X_0 определяется ее поведением в начальный момент времени $r=0$, т. е. в окрестности точки x_0 , и геометрией объемлющего многообразия. Эта оценка точна в том смысле, что существуют богатые серии троек (M, f, X) , для которых неравенство превращается в равенство. В этих случаях мы получаем точное выражение для функции объема и для объема минимальной поверхности.

Итак, для получения оценки на объем поверхности нужно сначала подсчитать на поверхности уровня $\{f=r\}$ модуль градиента функции f (относительно метрики, заданной на M), умножить его на коэффициент деформации поля $\operatorname{grad} f$, найти максимум α этого произведения на поверхности уровня и проинтегрировать функцию $1/\alpha$ по r . Конечно, функция $q(r)$, задаваемая неопределенным интегралом, определена с точностью до постоянного множителя, но это не влияет на оценку, так как в формулу для l входит $1/q(a)$, что устраняет неопределенность при умножении $q(r)$ на l .

Теорема 1 следует из общего утверждения о поведении нормированной функции объема. Пусть тройка (M, f, X) удовлетворяет условиям теоремы 1, тогда кусочно непрерывная функция $\psi(r)/q(r)$ не убывает по r при росте r от 0 до 1.

Отсюда получаем, что $\frac{\psi(1)}{q(1)} = \frac{\operatorname{vol}_k X_0}{q(1)} \geq \lim_{a \rightarrow 0} \frac{\psi(a)}{q(a)} = l$, т. е.

$\operatorname{vol}_k X_0 \geq l \cdot q(1)$. Аналогичное утверждение имеет место и для замкнутых минимальных поверхностей, реализующих нетривиальные (ко)циклы в объемлющем многообразии M [16].

Рассмотрим следствия теоремы 1, например, задачу об оценке снизу объема минимальной поверхности X^k , проходящей через центр стандартного евклидова шара D^n в $\mathbb{R}^n$. Среди этого класса поверхностей естественно выделены плоские k -мерные диски — сечения шара D^n k -мерными плоскостями, проходящими через центр. Хорошо известно, что во многих случаях (например, когда поверхность X^k комплексно-аналитическая в $\mathbb{R}^{2n} = \mathbb{C}^n$) объем ее части, расположенной внутри шара, оценивается снизу объемом плоского k -мерного диска.

Оказывается, верно общее утверждение, справедливое для любых глобально минимальных поверхностей.

Предложение 1. (см. [16]). Пусть X^k — глобально минимальная поверхность в $\mathbb{R}^n$, проходящая через центр шара D^n радиуса r и имеющая границу на границе шара (в смысле класса A^* , см. выше). Тогда $\text{vol}_k X \cap D^n \geq \Psi(0) r^k \gamma_k \geq r^k \gamma_k = \text{vol}_k D^n$, где $\Psi(0)$ — плотность поверхности X в точке 0 , γ_k — объем единичного евклидова шара, а $r^k \gamma_k$ — объем k -мерного евклидова шара радиуса r . Кроме того, $\Psi(0) \geq 1$, и равенство имеет место в том и только в том случае, когда точка 0 регулярна на поверхности X^k . Итак, объем любой глобально минимальной поверхности любой коразмерности, проходящей через центр евклидова шара и имеющей границу на границе шара, не меньше, чем объем стандартного центрального плоского сечения шара, т. е. чем объем плоского k -мерного диска.

Пусть M^n — замкнутое связное компактное риманово многообразие, x_0 — некоторая его точка. Известно, что M представляется в виде клеточного комплекса, содержащего только одну n -мерную клетку, гомеоморфную диску D^n , граница которого приклеена к подкомплексу C , $\dim C \leq n - 1$. Тогда множество $M \setminus (C \cup x_0)$ гомеоморфно открытому диску D^n с выброшенной точкой x_0 . Рассмотрим на M функцию f , гладкую на $M \setminus (C \cup x_0)$, имеющую ровно один минимум на M в точке x_0 (будем считать, что $f(x_0) = 0$), принимающую максимальное значение, равное 1, на C и не имеющую критических точек на открытом многообразии $M \setminus (C \cup x_0)$. Множество всех таких функций обозначим $F(x_0)$. Рассматривая на $D^n \setminus x_0$ поле $\text{grad } f$, где $f \in F(x_0)$, мы можем, следуя изложенной конструкции, определить функцию

$$q(r) = \exp \int_0^r \frac{dr}{\max_{x \in \{f=r\}} (\kappa_k | \text{grad } f |)},$$

зависящую от точки x_0 и от выбора функции f .

Определение 1. Функцией k -мерного F -дефекта многообразия M в точке x_0 назовем функцию $\gamma_k q(1) \lim_{a \rightarrow 0} \frac{a^k}{q(a)} = \Omega(x_0, f)$, где $q(1) = \lim_{r \rightarrow 1} q(r)$. В каждой точке $x_0 \in M$ рассмотрим число $\Omega_k(x_0) = \sup_{f \in F(x_0)} \Omega(x_0, f)$. Назовем, наконец, k -мерным F -дефектом многообразия M число $\Omega_k = \inf_{x_0 \in M} \Omega_k(x_0)$.

Это число зависит от размерности k и от M . Ясно, что всегда выполнено неравенство $\Omega_k > 0$. Рассмотрим (ко)гомологии $H_k^{(k)}(M, G)$ с коэффициентами в группе G . Пусть L — произвольная нетривиальная подгруппа в $H_k^{(k)}(M, G)$. Рассмотрим глобально минимальную поверхность X^k в M , реализующую эту подгруппу. Это означает, что при вложении $i: X \rightarrow M$ поверхность X реализует все (ко)циклы из подгруппы L . В слу-

чае гомологий должно выполняться условие $L \subset \text{Im } i_*$, где $i_*: H_k(X) \rightarrow H_k(M)$, а в случае когомологий: $L \subset H^k(M) \setminus \text{Ker } i^*$, где $i^*: H^k(M) \rightarrow H^k(X)$. Пусть $x_0 \in X$ и $\Psi(x_0)$ — значение функции плотности поверхности X в этой точке.

Теорема 2 (см. [16]). Пусть $X^k \subset M^n$ — замкнутая глобально минимальная поверхность в M , реализующая нетривиальную подгруппу L . Тогда $\text{vol}_k X \geq \Psi(x_0) \cdot \Omega_k \geq \Omega_k > 0$. Тем самым число Ω_k оказывается универсальной постоянной, оценивающей снизу k -мерный объем любой замкнутой минимальной поверхности, реализующей нетривиальный (ко)цикл в многообразии M . Существуют богатые серии примеров, когда оценка достигается на конкретных минимальных поверхностях нетривиального топологического типа, т. е. неравенство превращается в равенство. В частности, если для минимальной поверхности X выполнено равенство $\text{vol}_k X = \Omega_k$, то поверхность является гладким минимальным компактным замкнутым подмногообразием в M .

Эта теорема автора позволяет доказывать минимальность многих конкретных поверхностей. Например, если M — вещественное проективное пространство со стандартной метрикой, то все проективные подпространства, стандартно вложенные в M , являются глобально минимальными поверхностями, реализующими нетривиальные (ко)циклы. Аналогичное утверждение справедливо для комплексных и кватернионных подпространств в комплексном и кватернионном пространствах соответственно. Глобально минимальной поверхностью, реализующей нетривиальный (ко)цикл, является сфера S^8 в однородном пространстве $M^{16} = F_4/\text{Spin}(9)$ (стандартное вложение). Для всех этих поверхностей $\text{vol}_k X = \Omega_k$, и все они не только минимальные, но и вполне геодезические подмногообразия. Указанные поверхности имеют наименьший объем среди всех поверхностей, реализующих нетривиальные (ко)циклы в данной размерности. Если X — любое замкнутое кватернионное подмногообразие вещественной размерности 4 в кэлеровом кватернионном симметрическом связном компактном пространстве, то X — глобально минимальная поверхность в своем классе (ко)гомологий. Глобально минимальными поверхностями, реализующими нетривиальные (ко)циклы и обладающими свойством $\text{vol}_k X = \Omega_k$, являются также комплексные, кватернионные и вещественные грассмановы подмногообразия, стандартно вложенные в грассмановы многообразия того же типа, при некоторых ограничениях на размерность в вещественном случае. Другие примеры см. в [15, 16, 27, 33].

4. Случай отсутствия локальных минимумов функционала Дирихле в классе отображений однородных пространств в произвольное риманово многообразие.

Пусть M^m и N^n — гладкие римановы многообразия без края, причем M — компактно, замкнуто и ориентируемо. Мы будем интересоваться гармоническими отображениями $f: M \rightarrow$

$\rightarrow N$. С каждым гладким отображением $f: M \rightarrow N$ естественно связывается некоторое число $D[f]$, что позволяет определить многомерный функционал Дирихле. Пусть g_{ij} и $\hat{g}_{\alpha\beta}$ — римановы метрики на M и N соответственно. С отображением f естественно связан его дифференциал df , отображающий касательное пространство $T_x M$ в $T_{f(x)} N$. Пусть $x_1, \dots, x_m$ — локальные координаты на M^m , а $y_1, \dots, y_n$ — локальные координаты на N^n . Тогда отображение f записывается в виде системы функций $y_\alpha = y_\alpha(x_1, \dots, x_m)$, $1 \leq \alpha \leq n$. Пусть g^{ij} — элементы матрицы, обратной к матрице $\|g_{ij}\|$, т. е. $\sum_i g^{ij} g_{iq} = \delta_q^j$. Пусть

$d^m x$ — стандартная m -мерная внешняя форма риманова объема на M . Определим функционал Дирихле равенством $D[f] = \int_M \sum_{i,j,\alpha,\beta} g^{ij} \frac{\partial y_\alpha}{\partial x_i} \frac{\partial y_\beta}{\partial x_j} \hat{g}_{\alpha\beta} d^m x$. Таким образом, в вычислении интеграла Дирихле участвуют оба метрических тензора. Компактность и ориентируемость M предполагается для корректности определения.

Функционал Дирихле определен на бесконечномерном пространстве гладких (кусочно-гладких) отображений M в N . Иногда этот функционал называется также функционалом энергии. Гармонические отображения — это экстремали функционала Дирихле. В частном случае, когда M — двумерная поверхность, а $N = \mathbb{R}^3$, мы получаем гармонические поверхности, изученные в § 1 гл. 4, п. 12. Свойства функционала энергии и его экстремалей изложены, например, в [27]. Через $\delta_\eta D[f]$ будем обозначать первую вариацию (первую производную) функционала D в точке f по направлению векторного поля η , являющегося полем на многообразии N и определенном в окрестности образа $f(M)$ в N . Обычно рассматриваются поля определенного класса, например H_1^2 [1]. Пусть f_0 — критическая точка (экстремаль) функционала Дирихле, т. е. $\delta_\eta D[f_0] \equiv 0$ для любого возмущения (векторного поля) η в точке f_0 . Пусть $\delta^2 D$ — вторая вариация функционала Дирихле в точке f_0 . Удобно представлять ее как билинейную форму на линейном пространстве всех векторных полей, определенных в малой окрестности образа $f(M)$ в N . Рассмотрим максимальное подпространство векторных полей, на котором форма $\delta^2 D$ отрицательно определена. Векторные поля из этого подпространства характеризуются тем, что значение функционала уменьшается при возмущении функционала Дирихле вдоль них. Размерность этого максимального подпространства называется индексом критической точки f_0 , т. е. индексом гармонического отображения f_0 . Обозначим его через $\text{ind } f_0$. Если индекс гармонического отображения равен нулю, то отображение является локальным минимумом для функционала Дирихле. Если индекс положителен, то экстремаль f_0 является седловой точкой, т. е. существуют направления возмущения, вдоль ко-

торых энергия уменьшается. Известно [37], что индекс гармонического отображения всегда конечен. Чтобы доказать, что функционал Дирихле не имеет локальных минимумов в классе отображений $f: M \rightarrow N$, достаточно показать, что индексы всех гармонических отображений из этого класса строго положительны. Оказывается, существует много примеров многообразий M и N таких, что функционал Дирихле не имеет никаких локальных минимумов в классе отображений $f: M \rightarrow N$, отличных от тривиальных, т. е. от таких, когда M отображается в одну точку. Для описания таких примеров напомним некоторые определения.

Определение 2. Компактным неприводимым однородным пространством (многообразием) M мы будем называть однородное пространство $M = G/H$, полученное факторизацией компактной группы Ли G по компактной подгруппе H , причем линейное представление подгруппы H на касательном пространстве к M в точке $x_0 = eH$ является неприводимым, т. е. не имеет нетривиальных инвариантных подпространств.

Пространство M отождествляется с множеством левых классов смежности gH по подгруппе H . Она является стационарной подгруппой точки $x_0 = eH$, т. е. оставляет ее на месте при естественном правом действии H на M , когда $gH \rightarrow hgH$, где $g \in G$, $h \in H$. Так как точка x_0 неподвижна, то касательное пространство к M отображается в себя дифференциалом каждого преобразования из H . Это и задает линейное представление H на $T_{x_0}M$. При этом мы считаем, что на группе G задана биинвариантная римановая метрика (существующая на любой компактной группе Ли), которая естественно индуцирует некоторую метрику на M , инвариантную при действии G . Следовательно, G действует на M изометриями. Например, неприводимым компактным однородным пространством является сфера S^{n-1} со стандартной метрикой, допускающая представление в виде $SO(n)/SO(n-1)$, где подгруппа $SO(n-1)$ стандартно вложена в $SO(n)$.

Пусть $M = N$ и $f: M \rightarrow N$ — тождественное отображение. Легко видеть, что оно является гармоническим. Следовательно, определен индекс тождественного отображения M на себя. Сформулируем результаты, полученные А. И. Плужниковым.

Теорема 3. Пусть M — компактное неприводимое однородное риманово пространство и $f: M \rightarrow N$ — непостоянное (т. е. отличное от отображения в точку) гармоническое отображение M в любое гладкое риманово многообразие N . Предположим, что индекс тождественного отображения многообразия M на себя положителен. Тогда индекс отображения $f: M \rightarrow N$ также положителен. Другими словами, среди гармонических отображений M в N , отличных от отображения в точку, нет локальных минимумов функционала Дирихле.

Прежде чем привести конкретные примеры таких многообразий, сделаем комментарий. Результаты такого типа указы-

вают на глубокие различия между функционалом Дирихле и функционалом объема поверхности. Как мы знаем, в каждом нетривиальном гомотопическом классе фиксированного отображения $f: M \rightarrow N$ существует отображение f_0 , гомотопное исходному, на котором функционал объема достигает абсолютного минимума (см. гл. 4, § 2). Более того, всегда существует абсолютный минимум в любом нетривиальном классе бордизмов (гл. 4, § 2). В отличие от этих результатов, как мы видим, существуют замкнутые многообразия M , для которых в классе отображений $f: M \rightarrow N$, гомотопных непостоянному отображению, нет локальных минимумов, т. е. функционал Дирихле не только не имеет локальных минимумов, но и, в частности, не достигает абсолютного минимума на классе гомотопически нетривиальных (нестягиваемых) отображений. Здесь мы имеем дело с замкнутым многообразием M . Если же рассматривать многообразия с краем, то известны случаи, когда функционал Дирихле достигает абсолютного минимума в классе гомотопически нетривиальных отображений (см., например, [27]). Укажем теперь примеры многообразий, удовлетворяющих условиям теоремы.

Предложение 2. Никакое непостоянное гармоническое отображение следующих компактных замкнутых римановых многообразий M в любое риманово многообразие N не является локальным минимумом функционала Дирихле: 1) стандартные сферы S^m при $m \geq 3$, 2) кватернионные грассмановы многообразия $Sp(p+q)/(Sp(p) \times Sp(q))$, 3) многообразия $SU(2p)/Sp(p)$ при $p > 1$.

Все эти многообразия являются компактными неприводимыми однородными пространствами, для которых индекс тождественного отображения положителен [38]. Если многообразие M является сферой, то утверждение предложения 2 может быть усилено. Во-первых, удастся получить более сильную оценку на индекс произвольного гармонического отображения, а во-вторых, можно показать, что недостижимый абсолютный минимум функционала Дирихле равен нулю.

Предложение 3. Пусть $f: S^m \rightarrow N$ — непостоянное гармоническое отображение стандартной сферы, где $m \geq 3$, в произвольное гладкое риманово многообразие N . Тогда выполнено неравенство $\text{ind } f \geq m + 1$.

Эта оценка в общем случае неулучшаема, так как в [38] доказано, что в частном случае тождественного отображения сферы на себя (при $m \geq 3$) его индекс в точности равен $m + 1$.

Предложение 4. Абсолютный минимум функционала Дирихле на любой компоненте связности пространства $C^\infty(S^m, N)$, гладких отображений стандартной сферы S^m при $m \geq 3$ в произвольное риманово многообразие N равен нулю. Другими словами, в любом гомотопическом классе отображений сферы S^m в N всегда есть отображения со сколь угодно малой положительной энергией.

Более того, если многообразие M таково, что абсолютный минимум функционала Дирихле равен нулю на классе всех отображений, гомотопных тождественному отображению M на себя, то минимум функционала Дирихле равен нулю и на любом гомотопическом классе отображений M в произвольное гладкое риманово многообразие N .

5. Случай отсутствия локальных минимумов функционала Дирихле в классе отображений произвольного риманова многообразия в однородное пространство.

Оказывается, существуют факты, в некотором смысле двойственные результатам предыдущего пункта. Рассмотрим гладкие отображения произвольного гладкого риманова многообразия M в компактное неприводимое однородное пространство N . Другими словами, поменяем ролями многообразия M и N из предыдущего пункта, сохранив, впрочем, прежние обозначения для образа и прообраза. Приведем результаты, полученные А. В. Тыриным.

Теорема 4. Пусть M — связное компактное ориентируемое замкнутое гладкое риманово многообразие, а N — компактное неприводимое однородное пространство, снабженное естественной инвариантной римановой метрикой. Предположим, что индекс тождественного отображения многообразия N на себя положителен. Тогда любое гармоническое отображение $f: M \rightarrow N$, не являющееся постоянным (т. е. отличное от отображения в точку), также имеет положительный индекс, т. е. не является точкой локального минимума для функционала Дирихле.

Таким образом, как и в п. 4, вопрос о положительности индекса тождественного отображения неприводимого однородного пространства на себя является центральным при решении задачи о существовании локальных минимумов функционала энергии на пространствах отображений $f: M \rightarrow N$. Вычисление индекса тождественного отображения для многих случаев проведено в [38] (Смит).

Предложение 5 (см. [38]). Пусть $N = G/H$ — компактное неприводимое симметрическое пространство, где G — одна из классических групп Ли (т. е. G не является особой группой). Тогда индекс тождественного отображения N на себя всегда равен нулю, за исключением следующих случаев:

- 1) сферы S^m при $m \geq 3$;
- 2) $Sp(p+q)/(Sp(p) \times Sp(q))$;
- 3) $SU(2p)/Sp(p)$ при $p \geq 1$.

Комбинируя этот результат с теоремой 4, получаем утверждение:

Предложение 6. Из всех классических компактных неприводимых симметрических пространств $N = G/H$ для пространств: S^n , $n \geq 3$; $Sp(p+q)/(Sp(p) \times Sp(q))$; $SU(2p)/Sp(p)$, $p > 1$, и только для них выполнено следующее свойство: никакое гладкое непостоянное отображение $f: M \rightarrow N$ любого связ-

ного компактного ориентируемого риманова многообразия M в многообразии N не является точкой локального минимума для функционала Дирихле. В частности, в этих случаях не достигается и абсолютный минимум энергии.

Для всех перечисленных многообразий N можно указать тот класс деформаций, среди которых для любого гармонического отображения $f: M \rightarrow N$ заведомо найдется деформация, уменьшающая значение функционала Дирихле на этом отображении. Для каждого гармонического отображения f существует, вообще говоря, своя сокращающая деформация из этого класса.

Рассмотрим частный случай, когда $N = S^n$, $n \geq 3$. В этом случае можно явно предъявить класс деформаций, содержащий сокращающую деформацию, т. е. уменьшающую значение функционала Дирихле для гармонического отображения $f: M \rightarrow S^n$, т. е. дающую ненулевой вклад в индекс критической точки f . Для этого достаточно рассмотреть ограничения на сферу S^n (стандартно вложенную в $\mathbb{R}^{n+1}$) координатных линейных функций, порожденных декартовыми координатами в $\mathbb{R}^{n+1}$. Другими словами, рассмотрим на сфере S^n функции $\lambda_i(x) = x_i$, где $1 \leq i \leq n+1$, x_i — i -я декартова координата точки x . Наконец, возьмем градиенты этих функций. В результате получим набор из $n+1$ векторных полей на сфере. Эти поля и задают класс деформаций, содержащих деформацию, уменьшающую значение функционала Дирихле в критической точке, являющейся тождественным отображением.

В перечисленных результатах не фигурирует двумерная сфера. Дело в том, что для этого многообразия утверждения теорем 3 и 4 неверны. Известно (см. [37]), что любое голоморфное отображение кэлеровых многообразий реализует минимум функционала Дирихле в своем гомотопическом классе. Поэтому, в частности, тождественное отображение S^2 на себя является гармоническим, реализующим минимум функционала Дирихле в своем гомотопическом классе. Элементарное доказательство этого последнего факта для S^2 см., например, в [1].

В [39] была построена серия гармонических отображений сфер $S^m \rightarrow S^n$, однако вопрос о том, являются ли они локальными минимумами функционала Дирихле, оставался открытым. Из предложения 6 следует, что эти отображения не являются локальными минимумами энергии при $n \geq 3$. Интересно продолжить вычисление индексов тождественных отображений для компактных неприводимых пространств, отличных от рассмотренных в [38]. Изучим случай, когда таким пространством является простая компактная группа Ли, т. е. или классическая группа $A_l = \text{SU}(l+1)$, $l \geq 1$; $B_l = \text{SO}(2l+1)$, $l \geq 2$; $C_l = \text{Sp}(2l)$, $l \geq 2$; или $D_l = \text{SO}(2l)$, $l \geq 3$; или же одна из особых групп G_2, F_4, E_6, E_7, E_8 (А. В. Тырин).

Предложение 7. Пусть $N = G$ — простая компактная группа Ли A_l , $l \geq 1$; B_l , $l \geq 2$; C_l , $l \geq 2$; D_l , $l \geq 3$; G_2, F_4, E_6, E_7, E_8

с биинвариантной метрикой. Тогда индекс тождественного отображения N на себя равен нулю, за исключением следующих случаев: 1) C_l при $l \geq 2$; 2) D_3 ; 3) A_l , $l \geq 1$.

Используя теорему 4, получаем следующее утверждение:

Предложение 8. Из всех простых односвязных компактных групп Ли $N=G$ для групп $SU(l+1)$, $l \geq 1$; $Sp(2l)$, $l \geq 2$ и $Spin(6)$ и только для них никакое гладкое непостоянное отображение $f: M \rightarrow N$ любого связного компактного ориентируемого риманова многообразия M в группу N не является точкой локального минимума для функционала Дирихле.

Перечисленные результаты касаются однородных пространств, снабженных естественными инвариантными метриками. Возникает вопрос об устойчивости этих результатов относительно малых возмущений римановых метрик. Оказывается, в некоторых случаях устойчивость имеется (А. В. Тырин).

Пусть N^n — компактное замкнутое гладкое многообразие в сфере S^p , причем N не лежит ни в каком экваторе S^{p-1} , получающемся сечением сферы S^p гиперплоскостью, проходящей через центр. Пусть B — вторая фундаментальная форма вложения подмногообразия $N \subset S^p$. Определение и свойства формы B см., например, в [1, 27]. Эта форма $B(X, Y)$ определена на парах векторов X, Y , касательных к подмногообразию N . Пусть $e_1, \dots, e_n$ — ортонормированный базис в касательной плоскости к N . Тогда норму $\|B\|$ формы B определим так:

$$\|B\|^2 = \sum_{i,j=1}^n |B(e_i, e_j)|^2, \text{ где } |B(X, Y)| \text{ обозначает длину вектора}$$

$B(X, Y)$. Напомним, что форма B принимает значения в множестве векторов, ортогональных к N .

Теорема 5. Пусть N^n — компактное гладкое замкнутое подмногообразие в стандартной сфере S^p и пусть B — вторая фундаментальная форма вложения N в S^p . Предположим, что вложение $N \subset S^p$ удовлетворяет условию:

$$\left(1 + \max \left\{1, \frac{\sqrt{n-1}}{2}\right\}\right) \cdot \|B\|^2 < n - 2,$$

где $n = \dim N \geq 3$. Тогда для любого гладкого непостоянного гармонического отображения f произвольного гладкого связного ориентируемого компактного замкнутого многообразия M в многообразие N индекс этого отображения положителен. В частности, никакое непостоянное гармоническое отображение $f: M \rightarrow N$ не является точкой локального минимума для функционала Дирихле, и он вообще не достигает абсолютного минимума на любом нетривиальном гомотопическом классе отображений $f: M \rightarrow N$.

Итак, предложение 6, справедливое для стандартной сферы S^n , $n \geq 3$, остается справедливым и для всех римановых метрик на ней, получающихся малым возмущением стандартной метрики. Условие теоремы 5 означает, что подмногообразие N «не

очень сильно искривлено» в сфере S^p , так как ограничение на норму второй формы означает, что N не делает резких локальных поворотов в направлении своих нормальных плоскостей. Если N — плоское сечение сферы, то форма B равна нулю и условие теоремы 5 выполнено. Запас многообразий N , для которых выполнено неравенство теоремы 5, невелик. Такое многообразие является гомотопической сферой и, следовательно, при $n \geq 4$ гомеоморфно стандартной сфере. Поэтому теорему 5 можно рассматривать как теорему об устойчивости утверждения об отсутствии локальных минимумов энергии в классе отображений $f: M \rightarrow S^n$ при достаточно малом возмущении метрики на стандартной сфере S^n .

В заключение приведем теорему А. И. Плужникова.

Теорема 6. Пусть M — гладкое связное замкнутое ориентированное многообразие. Тогда следующие утверждения эквивалентны.

1) Многообразие M двухсвязно, т. е. первые его две гомотопические группы равны нулю ($\pi_1(M) = \pi_2(M) = 0$).

2) Хотя бы в одной римановой метрике на многообразии M минимум функционала энергии равен нулю на гомотопическом классе тождественного отображения многообразия M в себя.

3) Для любого гладкого риманова многообразия N и при произвольном выборе римановой метрики на многообразии M минимум функционала энергии равен нулю на всех компонентах связности пространства отображений $C^\infty(M, N)$.

4) Для любого гладкого компактного ориентированного риманова многообразия N и при произвольном выборе римановой метрики на многообразии M минимум функционала энергии равен нулю на всех компонентах связности пространства отображений $C^\infty(N, M)$.

5) Для любого гладкого компактного ориентированного риманова многообразия N и для любой римановой метрики на многообразии M глобальный минимум функционала энергии достигается лишь в одном гомотопическом классе отображений из N в M , а именно — на локально постоянных отображениях.

Существенно новым в этой теореме является условие двухсвязности многообразия, эквивалентное условиям 2—5. Другие утверждения теоремы опираются на известные результаты Сакса, Уленбек, Иллса, Смита, Сампсона, Leung'a.

ЛИТЕРАТУРА

1. Дубровин Б. А., Новиков С. П., Фоменко А. Т. Современная геометрия. — М.: Наука, 1979.
2. Милнор Дж. Теория Морса. — М.: Мир, 1965.
3. Милнор Дж. Теорема об h -кобордизме. — М.: Мир, 1969.
4. Фукс Д. Б., Фоменко А. Т., Гутенмахер В. Л. Гомотопическая топология. — М.: Изд-во Моск. ун-та, 1969.
5. Мищенко А. С., Фоменко А. Т. Курс дифференциальной геометрии и топологии. — М., Изд-во Моск. ун-та, 1980.
6. Рашевский П. К. Риманова геометрия и тензорный анализ. — М.: Наука, 1967.
7. Reifenberg E. R. Solution of the Plateau problem for m -dimensional surfaces of varying topological type. — *Acta Math.*, 1960, 104, N 1, 1—92.
8. Morrey Ch. B. Multiple integrals in the calculus of variations. — Berlin; Springer, Bd. 130, 1966.
9. Federer H. Geometric measure theorie. — Berlin; Springer, Bd. 153, 1969.
10. Almgren F. J. Existence and regularity almost everywhere of solutions to elliptic variational problem among surfaces of varying topological type and singularity structure. — *Ann. Math.*, Ser. 2, 1968, 87, N 2, 321—391.
11. Almgren F. J., Taylor J. E. The geometry of soap films and soap bubbles. — *Scientific American*, 1976, July, 82—93.
12. Фоменко А. Т. Многомерные вариационные методы в топологии экстремалей. — *УМН*, 1981, 36, № 6, 105—135.
13. Фоменко А. Т. Многомерная задача Плато в римановых многообразиях. — *Матем. сб.*, 1972, 89, № 3, 475—520.
14. Фоменко А. Т. Минимальные компакты в римановых многообразиях и гипотеза Райфенберга. — *Известия АН СССР*, 1972, 36, № 5, 1049—1080.
15. Фоменко А. Т. Многомерные задачи Плато на римановых многообразиях и экстраординарные теории гомологий и когомологий. Часть I. — *Труды семинара по векторному и тензорному анализу*, 1974, 17, 3—176; часть II. — *Труды семинара по векторному и тензорному анализу*, 1978, 18, 4—93.
16. Фоменко А. Т. О минимальных объемах топологических глобально минимальных поверхностей в кобордизмах. — *Известия АН СССР*, 1981, 45, № 1, 187—212.
17. Арнольд В. И., Варченко А. Н., Гусейн-Заде С. М. Особенности дифференцируемых отображений. — М.: Наука, 1982.
18. Nitsche J. C. Vorlesungen über Minimalflächen. — *Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen*, — Berlin; Springer, Bd. 199, 1975.
19. Poston Tim. The Plateau problem. An invitation to the whole of mathematics, Summer College on global analysis and its applications, 4 July — 25 August 1972. — International Centre for theoretical physics. Trieste, Italy, 1972.
20. Гуревич В., Волман. Теория размерности. — М.: ИЛ, 1948.
21. Борисович Ю. Г., Близняков Н. М., Израилевич Я. А., Фоменко Т. Н. Введение в топологию. — Москва: Высшая школа, 1980.
22. Bombieri E., De Giorgi, Giusti E. Minimal cones and the Bernstein problem. — *Invent. Math.*, 1969, 7, N 3, 243—268.

23. Simons J. Minimal varieties in Riemannian manifolds. — *Ann. Math.*, 1968, 88, N 1, 62—105.
24. Целочисленные потоки и минимальные поверхности / Сборник переводов. — М.: Мир, 1973.
25. Hsiang W. Y., Lawson H. B. Minimal submanifolds of low cohomogeneity. — *J. Diff. Geometry*, 1971, 5, N 1, 1—38.
26. Lawson H. B. The equivariant Plateau problem and interior regularity. — *Trans. Amer. Math. Soc.*, 1972, 173, 231—249.
27. Фоменко А. Т. Вариационные методы в топологии. — М.: Наука, 1982.
28. Дао Чонг Тхи. Мультиварифолды и классические многомерные задачи Плато. — *Известия АН СССР*, 1980, 44, № 5, 1031—1065.
29. Plateau J. A. F. *Statique Experimentale et Théorique des Liquides Soumis aux Seules Forces Moléculaires*. — Paris: Gauthier — Villars, 1873.
30. Lawson H. B., Osserman R. Non-existence, non-uniqueness and irregularity of solutions to the minimal surface system. — *Acta Math.*, 1977, 139, N 1.
31. Brothers J. E. Invariance of solutions to invariant parametric variational problems. — *Trans. Amer. Math. Soc.*, 1980, 262, N 1.
32. Rado T. On the problem of Plateau. — *Ergebnisse der Mathematik und ihrer Grenzgebiete*. — Berlin: Springer, 1933.
33. Дао Чонг Тхи. Многомерная вариационная задача в симметрических пространствах. — *Функц. анализ*, 1978, 12, № 1, 72—73.
34. Фоменко А. Т. Дифференциальная геометрия и топология. Дополнительные главы. — М.: Изд-во Моск. ун-та, 1983.
35. Suzuki S. On homeomorphisms of a 3-dimensional handlebody. — *Canadian J. Math.*, 1977, 29, N 1, 111—124.
36. Володин И. А., Фоменко А. Т. Об одном топологическом алгоритме. — В кн.: Тезисы докладов VI Всесоюзной топологической конференции. — Минск, Институт математики АН БССР, 1977, с. 44.
37. Eells J., Lemaire L. A report on harmonic maps. — *Bull. London Math. Soc.*, 1978, 10, N 1, 1—68.
38. Smith R. T. The second variation formula for harmonic mappings. — *Proc. Amer. Math. Soc.*, 1975, 47, N 1, 229—236.
39. Smith R. T. Harmonic mappings of spheres. — *Amer. J. Math.*, 1975, 97, N 2, 364—387.
40. Володин И. А., Кузнецов В. Е., Фоменко А. Т. О проблеме алгоритмического распознавания стандартной трехмерной сферы. — *УМН*, 1974, 24, № 5, 71—168.
41. Fomenko A. T. Multidimensional Plateau problem on Riemannian manifolds. On the problem of the algorithmical recognizability of the standart three-dimensional sphere. — In: *Proc. of the Intern. Congress of Math.* 1974, 515—525. — Vancouver, 1974.
42. Whitehead J. H. C. On certain sets of elements in a free group. — *Proc. London Math. Soc.*, Ser. 2, 1936, 41, 48—56.
43. Waldhausen F. Heegard—Zerlegungen der 3-Sphere. — *Topology*, 1968, 7: 2, 195—203.
44. Haken W. Theorie der Normalflächen. — *Acta Math.*, 1961, 105, 245—375.
45. Birman J. S., Hilden H. M. Heegard splittings of branched coverings of S^2 . — *Trans. Amer. Math. Soc.*, 1975, 213, 315—352.
46. Homma T., Ochiai M., Takahashi M. An algorithm for recognizing S^3 in 3-manifolds with Heegard splittings of genus two. — *Osaka J. Math.*, 1980, 17, 625—648.
47. Виро О. Я., Кобельский В. Л. Гипотеза Володина — Кузнецова — Фоменко о диаграммах Хегора трехмерной сферы не верна. — *УМН*, 1977, 32, № 5, 175—176.
48. Володин И. А., Фоменко А. Т. Многообразия. Узлы. Алгоритмы. — Труды семинара по векторному и тензорному анализу, 1978, 18, 94—128.
49. Ochiai M. A counterexample to a conjecture of Whitehead and Volodin — Kuznetsov — Fomenko. — *J. Math. Soc. Japan.*, 1979, 31, 687—691.
50. Morikawa O. A counterexample in the case genus 3. — *Math. Sem. Notes of Kobe Univ.*, 1980, 8, N 2, 295—298.

51. Smale S. On the structure of manifolds. — Amer. J. Math., 1962, 84, N 3, 387—399.
52. Богоявленский О. И. О точной функции на многообразиях. — Мат. заметки, 1970, 8, № 1, 77—83.
53. Шарко В. В. О гладких функциях на многообразиях. — Препринт. Ин-т математики АН УССР. — Киев, 1979.
54. Рурк К., Сандерсон Б. Введение в кусочно-линейную топологию. — М.: Мир, 1974.
55. Кёрби Р., Шарлеман М. Восемь ликов гомологической трехмерной сферы Пуанкаре. — УМН, 1982, 37, № 5.
56. Séminaire N. Bourbaki, vol. 1981—1982. Exposé 585—590, Fevrier, 1982.
57. Штанько М. А. Решение задачи Менгера в классе компактов. — ДАН СССР, 1971, 201, № 6, 1299—1302.
58. Чернавский А. В. О k -стабильности гомеоморфизмов и объединении клеток. — ДАН СССР, 1968, 180, № 5, 1045—1047.
59. Новиков С. П. Многозначные функции и функционалы. Аналог теории Морса. — ДАН СССР, 1981, 260, № 1, 31—35.
60. Новиков С. П. Гамильтонов формализм и многозначный аналог теории Морса. — УМН, 1982, 37, № 5, 3—49.
61. Новиков С. П., Шмельцер И. Периодические решения уравнений Кирхгофа свободного движения твердого тела в идеальной жидкости и расширенная теория Люстерника — Шнирельмана — Морса (ЛМШ). I. — Функ. анализ, 1981, 15, № 3, 54—66; II. — Функц. анализ, 1981, 15, № 4.
62. Freedman M. H. The topology of four-dimensional manifolds. — J. of Diff. Geom., 1982, 17, N 3, 357—453.
63. Шарко В. В. Точные функции Морса на односвязных многообразиях с односвязным краем. — УМН, 1981, 36, № 5.
64. Yau S. S.-T. Kohn—Rossi cohomology and its applications to the complex Plateau problem. I. — Ann. of Math., 1981, 113, N 1, 67—110.
65. Federer H., Fleming W. H. Normal and integral currents. — Ann. Math., 1960, 72, 458—520.
66. Osserman R. Global properties of minimal surfaces in E^3 and E^n . — Ann. of Math., 1964, 80, N 2, 340—364.

62, 84,

Мат.

ринт.

ю. —

ерной

evrier,

ДАН

нении

еории

еории

чений

ти и

I. —

№ 4.

J. of

зиях

plex

Ann.

n. —

